Task Complexity and Performance in Individuals and Groups Without Communication

Aditya Gulati, Thuy Ngoc Nguyen, Cleotilde Gonzalez
Dynamic Decision Making Laboratory, Carnegie Mellon University
5000 Forbes Ave, Pittsburgh, PA 15213, USA
aditya.gulati@iiitb.org, ngocnt@cmu.edu, coty@cmu.edu

Abstract

While groups where members communicate with each other may perform better than groups without communication, there are multiple scenarios where communication between group members is not possible. Our work analyses the impact of task complexity on individuals and groups of different sizes while solving a goal-seeking navigation task without communication. Our major goal is to determine the effect of task complexity on performance and whether agents in a group are able to coordinate to perform the task more effectively despite the lack of communication. We developed a cognitive model of each individual agent that performs the task. We compare the performance of this agent with individual human performance, who worked on the same task. We observe that the cognitive agent is able to replicate the general behavioral trends observed in humans. Using this cognitive model, we generate groups of different sizes where individual agents work in the same goal-seeking task independently and without communication. First, we observe that increasing task complexity by design does not necessarily lead to worse performance in individuals and groups. We also observe that larger groups perform better than smaller groups and individuals alone. However, individual agents within a group perform worse than an agent working on the task alone. This effect is not the result of agents within a group covering less ground in the task compared to individuals alone. Rather, it is an effect resulting from the overlap of the agents within a group. Importantly, agents learn to reduce their overlap and improve their performance without explicit communication. These results can inform the design of AI agents in human-machine teams.

Introduction

While many people prefer working alone, some tasks are either too large or complicated to be taken on alone. Having a team of people is often crucial for success. However, a group is usually only effective if members in the group can work together towards their shared objectives. Communication is often seen as a vital tool for coordination between members of groups. Indeed, past studies have shown that groups where members can communicate with each other do better than those groups where such communication channels do not exist (King et al. 2011; Sumner and King 2011; Oesch and Dunbar 2018).

It is evident that communication is useful, and even though our tools for communication are better than they have ever been, there can be situations where members of a group cannot communicate with each other. Such situations can arise when people do not want to communicate to avoid being spotted (for example, a group of Navy SEALS raiding a building) or when people simply do not have access to communication systems (for example, a group of explorers split up in a network of underground caves). Thus, our work focuses on studying groups where members cannot communicate with each other. For this study, we use a search and rescue task in a simulated scenario called the Minimap (Nguyen and Gonzalez 2021), which is explained in detail in the following section.

There are multiple algorithms that focus on optimally solving search and rescue tasks (Becker, Blatt, and Szczerbicka 2013; Jensen 2013). However, an important goal in human-machine teaming is to create systems that can work well with humans and not just perform tasks optimally (Bansal et al. 2019). For this, it is important to understand how humans behave. While Reinforcement Learning (RL) (Sutton and Barto 2018; Gershman and Daw 2017) has been shown to capture some trends in human behaviour (Gureckis and Love 2009; Simon and Daw 2011) and is widely used, it is focused on finding optimal solutions and not on understanding how humans make decisions (Botvinick et al. 2019). Thus, we focus on using cognitive models to predict how a group of humans will behave on the Minimap.

To do this, we analyse data from simulations run with Instance-Based Learning Agents (IBL Agents), built based on Instance-Based Learning Theory (IBLT) (Gonzalez, Lerch, and Lebiere 2003). IBL models have been shown to model the human decision making process accurately, and they are useful tools to understand and predict human behaviour (Gonzalez and Dutt 2011, 2012; Dutt and Gonzalez 2012, 2015; Dutt, Ahn, and Gonzalez 2011). Our goal is to create IBL models of teams that do not communicate and use them to understand the advantages of working in a group and the impact it has on individual members in the group. In particular, we focus on how the complexity of the task plays a role in the performance of individuals and groups of different sizes.

We produce data for individual IBL agents performing the task, and for individuals and groups of sizes 3 and 6, where each group of agents performs the Minimap task in three different scenarios which vary in their degree of difficulty. We analyse the performance of human participants with individual agents, and the performance of agents within groups to determine how agents learn to coordinate and become more effective in solving the task without communication.

The Minimap Task

Gridworld tasks are often used in the study of AI as they provide a simple environment for agents to perform multiple tasks in a wide range of applications like navigation or search and rescue tasks. The simplicity of these tasks also makes them suitable for studies on various aspects of human behaviour and decision making (Nguyen and Gonzalez 2020b,a; Rabinowitz et al. 2018). While the simplicity of these tasks makes experiments easier to conduct and makes data collection easier, it is also important to understand if the behaviour we observe here scales up to more complex richer domains. Thus, building on past work that has created IBL models on gridworld tasks (Nguyen and Gonzalez 2020b,a; McDonald, Nguyen, and Gonzalez 2021), our focus here is on scaling up this work, using a more complex richer environment, that we call the Minimap (Nguyen and Gonzalez 2021).

The Minimap is a 50×100 grid which represents one floor of a building with multiple rooms which have caught fire. Potential victims are spread across the building and their injuries have different degrees of severity with some needing more urgent care than others. The goal of a participant is to rescue as many victims as possible in a stipulated time frame. Each cell on the Minimap can contain a wall, a victim or can be empty. During one run of the game, participants start from a predefined position and move around the empty slots on the grid in an attempt to find victims. The participants at any time have four possible actions - moving up, down, left, or right to the corresponding neighbouring cell.

Figure 1 presents the representations of the three search and rescue scenarios in the Minimap used in this study. Each scenario has two kinds of victims - the more severely injured *yellow* victims and the less severely injured *green* victims. All three scenarios have 24 green and 10 yellow victims. A participant walks around the empty cells (light grey) to search for these victims, but cannot walk through the walls which are represented by the dark grey cells in figure 1. The scenarios differ in the placement of victims and the number of obstacles on the map. These obstacles (encircled in red in figure 1) are walls which are placed in the middle of a path. These obstacles restrict a participant from taking the path they block, thereby forcing participants to search for longer paths to get around the obstacle. Thus, when there are more obstacles, the structural complexity of the task increases. Since each scenario has a different number of obstacles, they have been assigned 3 levels of complexity - *low* (2 obstacles), *medium* (4 obstacles) and *high* (6 obstacles).

IBL Models

IBL models are theoretically grounded cognitive models used to model human decisions from experience based on IBL theory (Gonzalez and Dutt 2011, 2012; Dutt and Gonzalez 2012, 2015; Dutt, Ahn, and Gonzalez 2011).

A key component of an IBL agent is its memory, where an agent stores its experiences. An agent gains experience in two ways: 1) Experiences can be *pre-populated* i.e., these are experiences the agent had before the task started. These are used to simulate prior knowledge that an agent has about the task. 2) Experiences can be *experienced in real time* i.e., the agent experiences new situations which add to its bank of knowledge while performing a particular task.

An experience in the memory is represented as an instance. An instance has three main parts, the situation, decision and utility. The situation is typically a set of attributes used to represent the current state of the environment. The attributes used to define the situation are typically observable features of the environment. The decision is the action the agent took when it was faced with the situation and the utility is the reward it received for taking that decision.

When an agent performing a task needs to decide, it looks for instances in its memory which are similar to the current situation and computes an activation function on each of them. The activation function represents how readily available an instance is in the memory (Anderson and Lebiere 2014). While, each attribute can have a different importance, our work here is based on the idea that every attribute that represents the situation is equally important. Thus, the following simplified version of the activation A_i is used for an instance i:

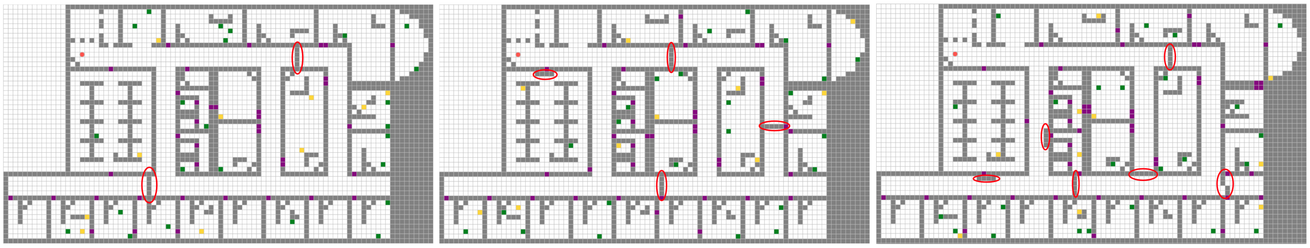
$$A_i = \ln\left(\sum_{t' \in \{1,\ldots,t-1\}} (t-t')^{-d}\right) + \sigma \ln\left(\frac{1-\gamma_i}{\gamma_i}\right) \quad (1)$$

where d is the decay and σ is the noise parameter. t' corresponds to every time step where the situation matched the current situation the agent is faced with and t is the current time step (this is used to capture the idea that it is harder to retrieve instances as they get older). The second part of the equation represents the noise and γ_i is a random number sampled from a uniform distribution $U(0,1)$.

Based on this activation function, the agent computes the expected utility for every possible action in the current situation. To do this, the agent uses a mechanism called *Blending* which combines the utility associated with all instances corresponding to situation s and action a. For this, it first calculates the retrieval probability of an instance as:

$$p_i = \frac{e^{A_i/\tau}}{\sum_{i=1}^{|l|} e^{A_i/\tau}} \quad (2)$$

where l is the set of all instances with situation s where action a was taken. τ is the temperature defined as $\sigma\sqrt{2}$. If the utility associated with the action a for instance i is u_i, then the blended value is computed as:



(a) Low (b) Medium (c) High

Figure 1: The three levels of the Minimap task. The light grey cells represent empty cells a participant can walk over. The dark grey cells are walls and the victims are represented by green and yellow cells. The obstacles have been encircled in red.

$$V(s,a) = \sum_{i=1}^{|l|} p_i u_i \tag{3}$$

The final action taken by the agent is the one with the highest blended value in the given situation.

IBL Agent For The Minimap Task

An IBL agent $\mathcal{A}_k$ makes decisions based on past experiences which are stored in memory in the form of triplets of the situation (s), the decision (d) and the utility (u). While our work builds upon the work of Nguyen and Gonzalez on the gridworld task (Nguyen and Gonzalez 2020a,b), the representation of an instance in memory was updated in order to deal with the additional complexities of the Minimap task.

The situation s represents the state of the agent in the environment and has two parts - 1) the location of the agent in the grid (the x-y coordinates) and 2) a bit vector to represent the victims rescued by $\mathcal{A}_k$ in the current episode. Each bit represents a victim and is set to 1 if the victim has been rescued in the current episode and 0 if it has not been rescued. The length of the bit vector is equal to the number of goals discovered by $\mathcal{A}_k$ across episodes i.e., a bit is added every time a victim is discovered by $\mathcal{A}_k$. Thus, at the start of every episode, all the bits are set to 0 and are set to 1 as and when the corresponding victim is rescued by $\mathcal{A}_k$.

In every situation, the agent needs to choose an action. For the Minimap task the agent has four possible actions - moving *up, down, left or right*. To make a decision in situation s at time t, the agent $\mathcal{A}_k$ computes the blended value of every possible action (equation 3) and picks the action with the highest utility. Thus, for every agent $\mathcal{A}_k$, in every episode, we can define a trajectory $\mathcal{T}_i = \{(s_t, d_t)\}_{t=0}^{T}$.

Each step in $\mathcal{T}_i$ is a part of $\mathcal{A}_k$'s experience on the task. However, to be stored as an instance it needs an associated utility. Since rewards are available only upon rescuing a victim steps in $\mathcal{T}_i$ are stored with a default utility temporarily. Once a victim is rescued by $\mathcal{A}_k$, the utility of all the instances corresponding to the steps in $\mathcal{T}_i$ which led to the victim are updated with the reward associated with rescuing the victim.

IBL models for Group Behaviour

IBL has also been used to model the behaviour of groups of humans (Lejarraga, Lejarraga, and Gonzalez 2014; Gonzalez et al. 2015; McDonald, Nguyen, and Gonzalez 2021). These studies have been performed on both static and dynamic environments but are performed often in simple choice tasks. This work expands on the past studies of IBL models for groups by testing the impact of group size and task complexity in the Minimap, a task with a large state space.

A group is modelled as multiple IBL Agents performing the task simultaneously without communicating with each other. While the environment is static for a single agent, performing the task with multiple agents makes the environment dynamic for each agent involved. For example, a victim $\mathcal{V}_j$ found by agent $\mathcal{A}_{||}$ in one episode can be rescued by agent $\mathcal{A}_{k'}$ in a subsequent episode before agent $\mathcal{A}_k$ reaches $\mathcal{V}_j$. Thus, for agent $\mathcal{A}_k$ the environment is dynamic making the task harder for agents in a group.

Experiment

We manipulated two factors to understand their impact on the performance of groups - the size of the group and the complexity of the task. The size of the the group involved 1 individual, 3, and 6; the task has three levels of complexity - low, medium, and high.

In total, we ran nine simulations, one for each group size (1, 3 or 6) on every level of task complexity (low, medium, and high). Each simulation was run for 50 identical agents (trials) and all results presented here have been averaged over these 50 trials.

Each group of agents ran for 50 episodes. Each episode was set up in the same way and had at most 2500 steps for each agent. An episode ended if the group rescued every victim or the limit of 2500 steps was reached. A victim could be rescued at most once in each episode. Once a victim was rescued, it "dissapeared" from the map until the start of the next episode. Rescuing a green victim gave the agent 0.25 points (r_{green}) while rescuing a yellow victim gave the agent 0.75 points (r_{yellow}) since the yellow victims have more serious injuries. If an agent tried to walk over a wall, it would stay in the same cell and receive a penalty of -0.05 points. If two agents tried to move to the same cell, one of them was chosen uniformly randomly and allowed to move while the other received a penalty of -0.01 points. The simulations were run on a machine with a 3.4GHz Intel(R) Core(TM) i7-4770 processor.

For every group, we measured the following parameters to analyse their behaviour:

- **Performance:** The performance (P) of an agent $\mathcal{A}_k$ is measured by the total reward collected for rescuing victims across N episodes i.e., if r_n is the reward collected by $\mathcal{A}_k$ in the n^{th} episode, then P is defined as:

$$P = \frac{\sum_{n=1}^{N} r_n}{\sum_{n=1}^{N} (N_{green} \times r_{green} + N_{yellow} \times r_{yellow})} \tag{4}$$

  The performance of a group is measured as the sum of the performances of every individual in the group.

- **Coverage:** The coverage is used to measure the ability of a group to explore the map in an episode. If L_k is the set of locations on the map visited by agent $\mathcal{A}_k$ and L_{map} is the set of all locations that an agent can visit, then the coverage of a group of size M is defined as:

$$\text{Coverage} = \frac{|\bigcup_{k=1}^{M} L_k|}{|L_{map}|} \tag{5}$$

- **Overlap:** The overlap measures in every episode the amount of common area explored by an agent $\mathcal{A}_k$ and the other agents in its group. It is defined as

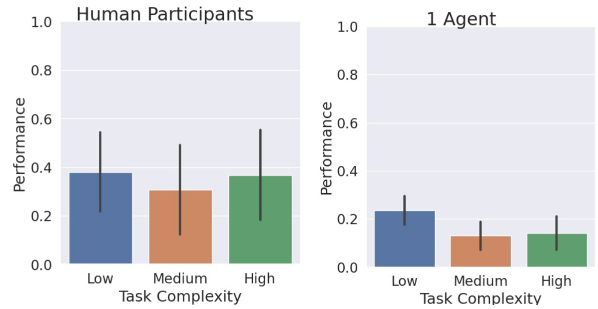
$$\text{Overlap} = \frac{|L_k \cap (\bigcup_{k' \neq k} L_{k'})|}{|L_k|} \tag{6}$$

- **Discovery Time:** This metric helps us understand how easy it is for a group to find victims. If V is the set of all victims rescued across all episodes, then the discovery time for a victim $v \in V$ is the first episode where the victim was rescued by any agent in the group. For a group of agents, the discovery time measured is the average discovery time of each of the rescued victims and ranges between 1 and 50.

Results

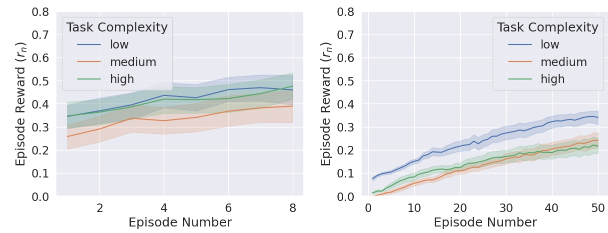
To understand the impact of group size and task complexity on the behaviour of groups, we start by looking at the performance of individual agents in relation to the performance of human participants in a data set in which individuals aim to do the Minimap task in an interactive experimental tool.

We compare the performance of independent IBL agents to human participants performing the Minimap task alone. This helps us establish whether the trends noted in the performance of individual IBL agents on the Minimap task are similar to those of human participants. This would support the expectation that predictions made in groups of IBL models may be observed in groups of humans as well.

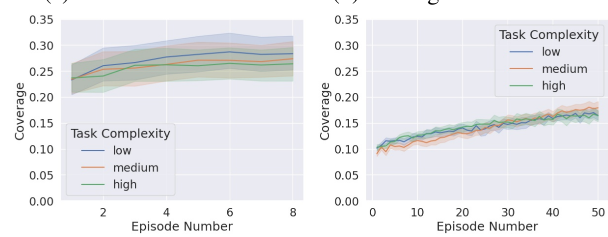


(a) Human Participants (b) IBL Agent

Figure 2: Performance of humans and agents on the Minimap task



(a) Human Performance (b) IBL Agent Performance



(c) Human Coverage (d) IBL Agent Coverage

Figure 3: Performance and coverage of human participants and a single IBL agent across episodes on the Minimap task

Humans and Individual IBL Agents

An experiment with an interactive version of the Minimap was conducted with human participants by Nguyen and Gonzalez and the data set has been made available publicly[1].

The data was collected from 297 participants performing the Minimap task under six different conditions. Out of these six conditions, three matched the situation of the Minimap task for the IBL agents in each of the three levels of complexity as described earlier. This resulted in a dataset of 149 participants distributed roughly equally between the three levels of structural complexity. Although not exactly equivalent to the task done by the IBL agents, human participants can be roughly compared to individual IBL agents as shown in Figure 2.

Figure 2a shows the average performance of the human participants and Figure 2b shows the performance of IBL agents performing the task independently. For each level of

[1] https://osf.io/5gmsc/?view_only=b7b13bcae1da448e8c3a5d58ad976e34

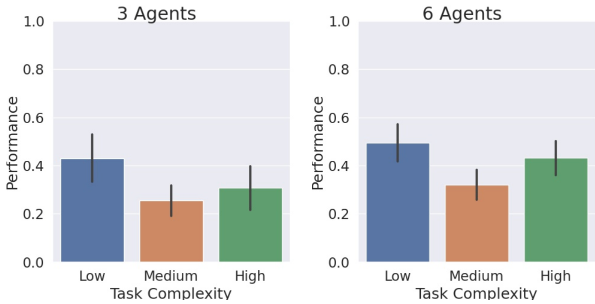


Figure 4: Relation between performance and task complexity for groups of size 3 and 6

complexity, 50 IBL agents were used to estimate performance. It is important to observe that the IBL agents were not fit to human data. The results presented here are pure predictions, based on the IBL theory. The IBL agents reflect how different complexity levels impact the performance of individual agents.

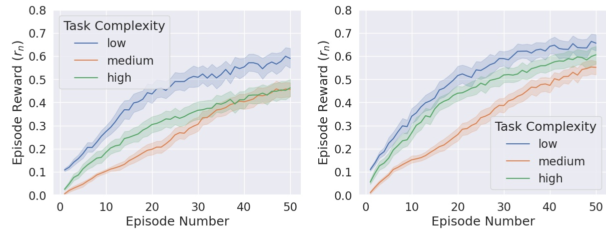
Figure 2 shows that humans as well as IBL agents perform worse on the medium structural complexity task compared to the other two levels of complexity. This contradicts the intuitive expected linear relation between complexity level and performance, suggesting that the design of task complexity that relies only on structural characteristics (i.e., the number of obstacles), does not necessarily result in a more complex task in terms of performance and decisions that humans or agents make.

To understand these trends better, we look at how human participants and IBL agents perform across episodes (figures 3a and 3b) and the coverage of human participants and IBL agents (figures 3c and 3d). It is clear that human participants outperform independent IBL agents. This is largely due to the fact that human participants are able to explore a larger portion of the grid. Additionally, we see that the complexity of the task has negligible impact on the ability of human participants to explore the map and similar trends can be seen for independent IBL agents.

The data from human participants was available only for individual participants and not for groups. The similarity in trends for individual human participants and independent IBL agents is encouraging. In the following sections, we focus on the performance of a group of IBL agents that do not communicate with each other and on the behaviour of individual agents within each group. The predictions of groups of IBL agents can be used as predictions about the possible behaviour of similarly structured groups of human participants.

Group Performance

Figure 4 shows the average performance with structural complexity for groups of 3 and 6 agents across three levels of structural complexity. Again, we observe that groups perform worse on the medium structural complexity compared to the other two levels of complexity, regardless of the group size. There appear to be a small advantage in larger groups,



(a) 3 Agents (b) 6 Agents

Figure 5: Trends of reward collected per episode for groups of size 3 and 6

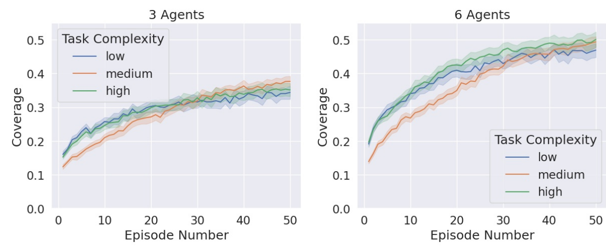


Figure 6: The area covered by a group over time for groups of size 3 and 6

where the performance is slightly better in the groups of 6 agents compared to groups of 3 agents. Additionally, the difference in performance between the medium and high complexity tasks also appears to be larger for larger in groups of larger size. To better understand these trends, we look at how the performance and coverage of groups changes across episodes.

Performance across episodes Figure 5 shows the variation in performance across episodes for groups of sizes 3 and 6, and for individual agents for all three levels of structural complexity. The x-axis represents the episode number and the y-axis represents the reward the group collected in the corresponding episode i.e., r_n from equation 4.

Figure 5a shows the reward collected over time for groups of three agents. Here again, the group is initially performing worse on the medium structural complexity task. However, as time passes, the agents learn to find victims, their performance starts picking up and comes close to the performance on the high structural complexity task. However, because there are more agents the effect seen with one agent are more pronounced i.e., the difference between performance on the medium structural complexity task and the other complexity levels is significantly higher. This effect is even more pronounced for larger groups as seen in Figure 5b. This indicates that it is not the number of obstacles that truly impact performance but rather how hard it is to find victims.

In line with these ideas, we computed the average time taken to find each victim in the different Minimap scenarios. We noted that on average, a victim was rescued for the first time around the 20^{th} episode by groups solving the medium complexity task in contrast to the low and high complexity

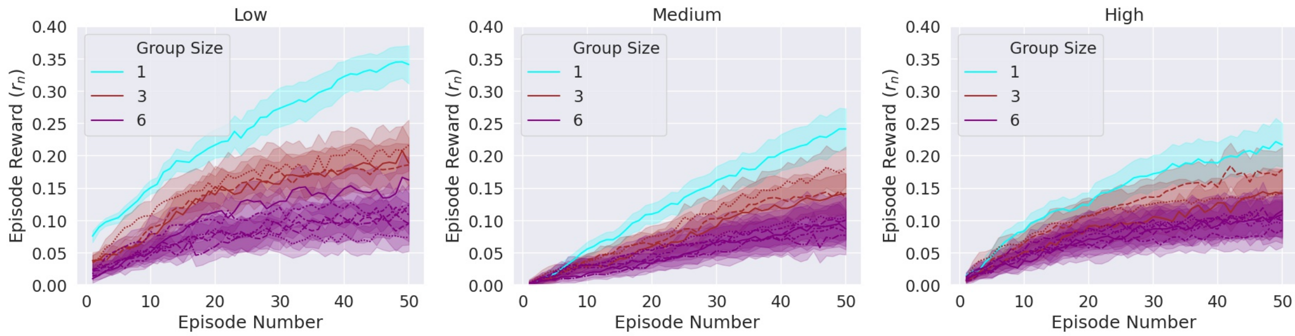


Figure 7: Performance of different agents in groups of size 3 and 6 for three levels of structural complexity.

task where victims were rescued for the first time around the 10^{th} episode. The average discovery time for victims stayed the same regardless of the size of the group. This independence of discovery time from group size coupled with the fact that the dip in performance on the medium complexity task was noted in human participants and independent IBL agents makes it likely that groups of humans will take longer to find victims on the medium complexity Minimap task compared to the low and high complexity task. This again, suggests that the structural complexity alone does not determine how complex the task can be for a group of agents.

Coverage On the Minimap task, the more ground a group can cover, the more likely the members of the group are to find victims. Intuitively, a larger group should be able to cover more ground, thereby allowing the group to rescue more victims.

Figure 6 shows the average group change in coverage over time for groups of size 3 and 6. The x-axis represents the episode number and the y-axis represents the coverage of an episode as defined in equation 5. As time passes, the groups are able to cover more ground in each episode. Additionally, larger groups cover more ground than smaller ones, which indicates that larger groups will perform better.

Individual Performance within Groups

While it is clear that larger groups perform better because they are able to cover more ground, it is also important to understand how individual agents in the group perform. Does the performance of a group increase because each agent performs better? Or does the performance of each agent stay the same and it is just more agents that allow groups to perform better? We answer these questions in this section.

Figure 7 shows individual agent performance in groups of 3 and 6 compared to when individual agents perform the task alone. For each plot, the x-axis represents the episode number and the y-axis represents the reward per episode i.e., r_n as used in equation 4. Each curve represents a single agent and the group size is indicated by the color.

The major observation is that individuals within a group perform worse than individuals that work in the task alone, regardless of the complexity of the task. Furthermore, the larger the group is, the worse the individual performance within a group is. To explain this effect, we look at how individuals within a group cover the task space (i.e., coverage) and how much they overlap with each other while doing the task.

Coverage While figure 6 revealed that larger groups cover more ground, figure 8 shows the coverage of each agent in a group and the coverage of independent IBL agents. While it is clear that each agent in a group covers less ground than an agent working alone, the difference in coverage is minor. This makes it clear that working in a group does not significantly hamper individual agents' ability to explore the map. However, this advantage is only effective if the agents are able to split up parts of the map effectively.

If individual agents are unable to split up effectively, larger groups may not be effective since some agents in these larger groups will just be repeating the work done by other agents. Splitting up may be easier for groups where agents can communicate (since they can plan out strategies to pick different areas) compared to groups like the ones studied here. Thus, we look at how effectively agents in groups are able to split up by looking at the overlap in coverage by agents in groups.

Overlap Figure 9 shows the overlap in every episode for every agent in a group of size 3 or 6 for all three levels of structural complexity. The x-axis indicates the episode number and the y-axis indicates the overlap measured in the corresponding episode as defined in equation 6.

While the overlap is understandably higher for agents in larger groups, it reduces over time regardless of task complexity. This indicates that the agents learn to find specific areas of the map and focus on them without getting in the way of their teammates - even without explicitly communicating with each other. The overlap for every agent decreases over time, but never goes down to 0 indicating that there is always some overlap between all members of a group. These trends explain how groups are able to cover a larger area over time and why the performance of an individual in a group stays below the performance of an agent acting alone.

Thus, each agent working in a group performs worse than an agent working alone, but groups as a whole still perform better than individuals and larger groups perform better than smaller ones.

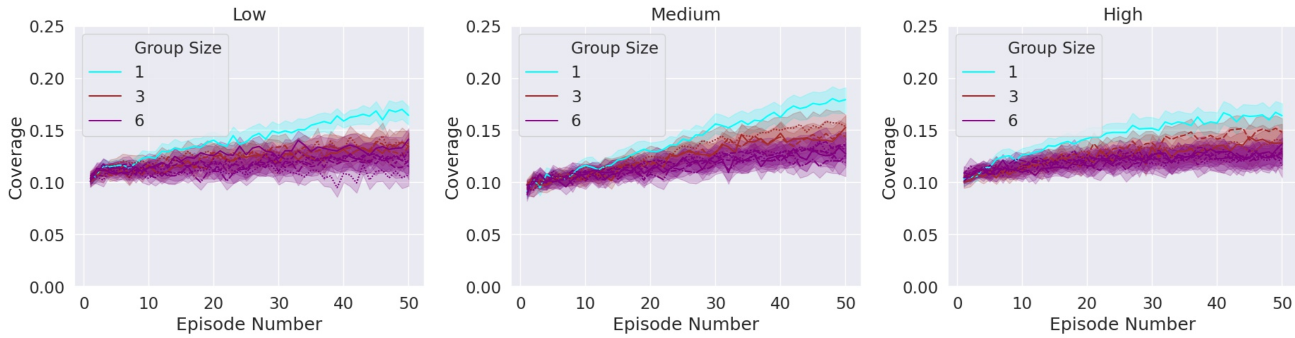


Figure 8: The area covered by a every member of a group over time for groups of size 3 and 6

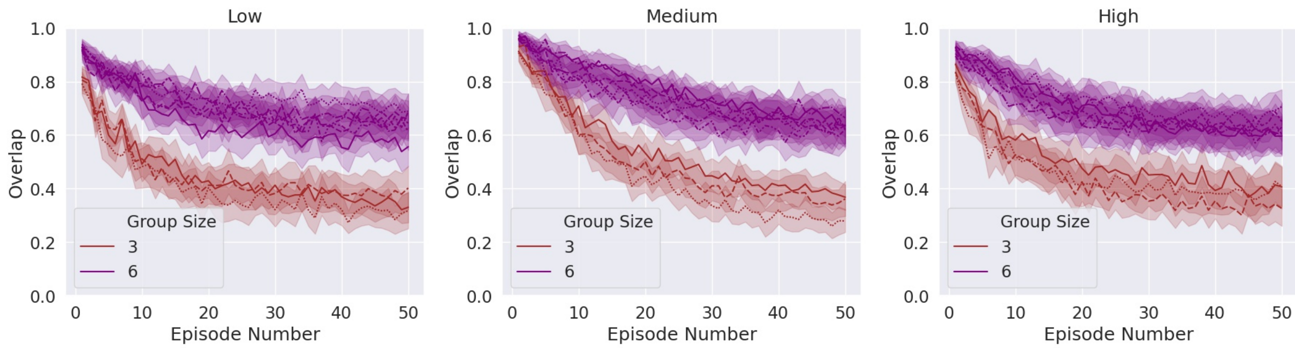


Figure 9: The overlap for every member of a group over time for groups of size 3 and 6

Conclusion

Working in groups is crucial and is something humans do often. To design AI that is able to work along with humans in groups, it is important to design models that emulate the way humans work in groups. Towards this end, we have worked on predicting the behaviour of humans in groups where members cannot communicate with each other. Particularly, we investigate how these groups of different sizes are impacted by task complexity. We created individual agents based on a cognitive theory of decisions from experience (Gonzalez, Lerch, and Lebiere 2003).

We saw that the trends for the performance of a single IBL agent are similar to those seen for human participants working on the task alone. We find that human participants as well as individual agents perform worse in the task of medium complexity than in the task of high complexity. This same effect is also observed on the average group performance regardless of the size of the group. This provides a lesson regarding how to design tasks of various complexities: defining complexity of a task based only on structural factors may not be enough to determine how complex a task will be in practice. We note that the reason the medium complexity map is harder, even though it has less obstacles than the high complexity map, is because it is harder to reach victims in this map.

In addition, we found that larger groups perform better, but the individual agents within a group perform worse than an agent attempting the same task alone. Moreover, the larger the group is, the worse individual agents within a group will perform. This effect does not seem to be due to the area that individual agents within a group cover compared to individual agents working alone. Rather, it seems that this is due to the overlap among agents. The overlap is greater in larger groups, but all agents within each group learn to improve their performance across episodes by reducing the amount of overlap between them - even without explicitly communicating.

Overall, we expect that these results will also hold for groups of human participants that do not communicate with each other and are important to consider while designing new AI for human-machine teams.

Acknowledgements

This research was sponsored by the Defense Advanced Research Projects Agency and was accomplished under Grant Number W911NF-20-1-0006.

References

Anderson, J. R.; and Lebiere, C. J. 2014. *The atomic components of thought*. Psychology Press.

Bansal, G.; Nushi, B.; Kamar, E.; Lasecki, W.; Weld, D.; and Horvitz, E. 2019. Beyond Accuracy: The Role of Mental Models in Human-AI Team Performance. In *HCOMP*. AAAI. URL https://www.microsoft.com/en-us/research/publication/beyond-accuracy-the-role-of-mental-models-in-human-ai-team-performance/.

Becker, M.; Blatt, F.; and Szczerbicka, H. 2013. A Multi-agent Flooding Algorithm for Search and Rescue Operations in Unknown Terrain. In *Multiagent System Technologies*, 19–28. Springer Berlin Heidelberg. doi:10.1007/978-3-642-40776-5_5. URL https://doi.org/10.1007/978-3-642-40776-5_5.

Botvinick, M.; Ritter, S.; Wang, J. X.; Kurth-Nelson, Z.; Blundell, C.; and Hassabis, D. 2019. Reinforcement Learning, Fast and Slow. *Trends in Cognitive Sciences* 23(5): 408–422. doi:10.1016/j.tics.2019.02.006. URL https://doi.org/10.1016/j.tics.2019.02.006.

Dutt, V.; Ahn, Y.-S.; and Gonzalez, C. 2011. Cyber situation awareness: Modeling the security analyst in a cyber-attack scenario through instance-based learning. In *IFIP Annual Conference on Data and Applications Security and Privacy*, 280–292. Springer.

Dutt, V.; and Gonzalez, C. 2012. The Role of Inertia in Modeling Decisions from Experience with Instance-Based Learning. *Frontiers in Psychology* 3: 177. ISSN 1664-1078. doi:10.3389/fpsyg.2012.00177. URL https://www.frontiersin.org/article/10.3389/fpsyg.2012.00177.

Dutt, V.; and Gonzalez, C. 2015. Accounting for outcome and process measures in dynamic decision-making tasks through model calibration. Technical report, Carnegie Mellon University Pittsburgh United States.

Gershman, S. J.; and Daw, N. D. 2017. Reinforcement Learning and Episodic Memory in Humans and Animals: An Integrative Framework. *Annual Review of Psychology* 68(1): 101–128. doi:10.1146/annurev-psych-122414-033625. URL https://doi.org/10.1146/annurev-psych-122414-033625.

Gonzalez, C.; Ben-Asher, N.; Martin, J. M.; and Dutt, V. 2015. A cognitive model of dynamic cooperation with varied interdependency information. *Cognitive science* 39(3): 457–495.

Gonzalez, C.; and Dutt, V. 2011. Instance-based learning: Integrating sampling and repeated decisions from experience. *Psychological review* 118(4): 523.

Gonzalez, C.; and Dutt, V. 2012. Refuting data aggregation arguments and how the instance-based learning model stands criticism: A reply to Hills and Hertwig (2012). *Psychological Review* 119(4): 893–898. doi:10.1037/a0029445. URL https://doi.org/10.1037/a0029445.

Gonzalez, C.; Lerch, J. F.; and Lebiere, C. 2003. Instance-based learning in dynamic decision making. *Cognitive Science* 27(4): 591–635. doi:https://doi.org/10.1207/s15516709cog2704_2. URL https://onlinelibrary.wiley.com/doi/abs/10.1207/s15516709cog2704_2.

Gureckis, T. M.; and Love, B. C. 2009. Short-term gains, long-term pains: How cues about state aid learning in dynamic environments. *Cognition* 113(3): 293–313. doi:10.1016/j.cognition.2009.03.013. URL https://doi.org/10.1016/j.cognition.2009.03.013.

Jensen, E. A. 2013. Dispersion and exploration for robot teams. In *Proceedings of the 2013 international conference on Autonomous agents and multi-agent systems*, 1437–1438.

King, A. J.; Narraway, C.; Hodgson, L.; Weatherill, A.; Sommer, V.; and Sumner, S. 2011. Performance of human groups in social foraging: the role of communication in consensus decision making. *Biology Letters* 7(2): 237–240. doi:10.1098/rsbl.2010.0808. URL https://royalsocietypublishing.org/doi/abs/10.1098/rsbl.2010.0808.

Lejarraga, T.; Lejarraga, J.; and Gonzalez, C. 2014. Decisions from experience: How groups and individuals adapt to change. *Memory & cognition* 42(8): 1384–1397.

McDonald, C.; Nguyen, T. N.; and Gonzalez, C. 2021. Multi-Agent Specialization and Coordination without Communication in a Gridworld Task. *ACM Collective Intelligence Conference* .

Nguyen, T. N.; and Gonzalez, C. 2020a. Cognitive Machine Theory of Mind. In *CogSci*.

Nguyen, T. N.; and Gonzalez, C. 2020b. Effects of Decision Complexity in Goal seeking Gridworlds: A Comparison of Instance Based Learning and Reinforcement Learning Agents. Technical report, Carnegie Mellon University.

Nguyen, T. N.; and Gonzalez, C. 2021. Minimap: A Dynamic Decision Making Interactive Tool for Search and Rescue Missions. Technical report, Carnegie Mellon University.

Oesch, N.; and Dunbar, R. I. M. 2018. Group size, communication, and familiarity effects in foraging human teams. *Ethology* 124(7): 483–495. doi:https://doi.org/10.1111/eth.12756. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/eth.12756.

Rabinowitz, N. C.; Perbet, F.; Song, H. F.; Zhang, C.; Eslami, S.; and Botvinick, M. 2018. Machine theory of mind. *arXiv preprint arXiv:1802.07740* .

Simon, D.; and Daw, N. 2011. Environmental statistics and the trade-off between model-based and TD learning in humans. In Shawe-Taylor, J.; Zemel, R.; Bartlett, P.; Pereira, F.; and Weinberger, K. Q., eds., *Advances in Neural Information Processing Systems*, volume 24. Curran Associates, Inc. URL https://proceedings.neurips.cc/paper/2011/file/c9e1074f5b3f9fc8ea15d152add07294-Paper.pdf.

Sumner, S.; and King, A. J. 2011. Actions speak louder than words in socially foraging human groups. *Communicative & Integrative Biology* 4(6): 755–757. doi:10.4161/cib.17701. URL https://doi.org/10.4161/cib.17701. PMID: 22446547.

Sutton, R. S.; and Barto, A. G. 2018. *Reinforcement learning: An introduction*. MIT press.