APPENDIX A
EXPERIMENT SETTING AND DETAILS

Considering cross-platform, scalability, and efficiency of the simulations, we chose the "Unity" game engine to simulate the *Explorers and Aliens Game* and selected Gambit [55] toolkit for calculating each level's Nash Equilibrium.

In the simulation, the Explorer agents follow the agent needs hierarchy principle to drive their behaviors, such as going back to base to charge, avoiding the collision, cooperation. So if the energy and HP amount are below a certain level which means the current safety and basic needs can not satisfy, they will stop against the opponents and leave the conflict area.

Furthermore, we assume Wi-Fi wireless device for communication, ultrasonic sensor for measuring the distance, camera for detecting the adversary and object, and laser sensor for measuring speed. And we also assume that agents use the particle beam weapon to attack each other in our simulation. However, with this assumption, we simulated the attack/defend process through the agent proximity and the time spent in the zone attacking/defending. The agent's sensing range is set to 3 m, and the communication range is 10 m for the partial scenarios. In the full cooperation scenarios, the communication range is 1000 m, which means the agents can always connect with each other.

Each interaction (trial) in the *Explorers and Aliens Game* lasts about 40 to 50 minutes in the simulated experiments on a laptop with Intel i7 Processor, GeForce GTX 1050 Ti GPU, and 16 GB DDR4 RAM running on Ubuntu 18.04. GUT can run the decision-making calculations in real-time. For instance, we provide the time taken to calculate the decisions at each level of the GUT in our simulations: For level one (2x2 actions - Table I) is 0.0288 seconds, level two (3x3 actions - Table II) is 0.0443 seconds and level three (2x3 actions - Table III) is 0.0311 seconds. In total, it took around 0.1s to make a strategic decision at a time. On the contrary, considering all strategy combinations in one game without any hierarchy (12x18 tactic matrix), it costs around 0.26s to arrive to the same decision. This 2.5x improvement in time efficiency with GUT is even more significant when the action space of the game is even larger. Also, for larger games, making this calculation in a distributed manner would further increase the efficiency.

In our experiments, we assume that explorers and aliens have the same moving speed, and aliens can not share information. It is worth noting that an alien attacking capability is three times that of an explorer in all the experiments, which means that aliens represent higher capabilities (who act based on their self-interests) in preventing explorers from their tasks.

Below, we clarify the terminologies on communication and cooperation variations used in the experiments.

Partial Communication An agent only communicates and shares information with other ally agents in their observable (sensing) range. They use this partial data in this modality.

Full Communication Agents always keep in touch with each other even not in the observable range, and the communication graph may resemble a completely connected graph. It allows each agent to communicate and exchange data with its neighbor until the group reaches *Information Equilibrium*, which means that every group member has the same information.

Algorithm 2: Obstacle avoidance algorithm to adapting the edge of the team formation shape.

Input: Explorers' and Obstacles' states
Output: moving direction r and distance Δd

while *The nearest collision point* c ! $= Null$ **do**
  calculate the number n and m of non-collision agents in both side of the line l passing through c and perpendicular c's tangent;
  if $n > m$ **then**
    r = n side in line l;
    Δd = one step of agent's movement;
  else if $n == m$ **then**
    agent stop;
  else if $n < m$ **then**
    r = m side in line l;
    Δd = one step of agent's movement;
return r = current position to goal point, r, Δd

Noncooperation (NC) Agents do not cooperate (or communicate) with others. They make decisions based on their own needs and benefits only, motivated by their *self-interest*.

Partial Cooperation (PC) Based on the *partially communicated* information, agents only cooperate with the observable group members to maximize their needs and minimize costs.

Full Cooperation (FC) According to the *fully communicated* data from all the agents in a group, agents make decision based on the *group-interest* showing *Collective Rationality*.

The video demonstration of the experiments showing sample trials of experiments using GUT and Greedy/QMIX is available at http://hero.uga.edu/research/gut/.

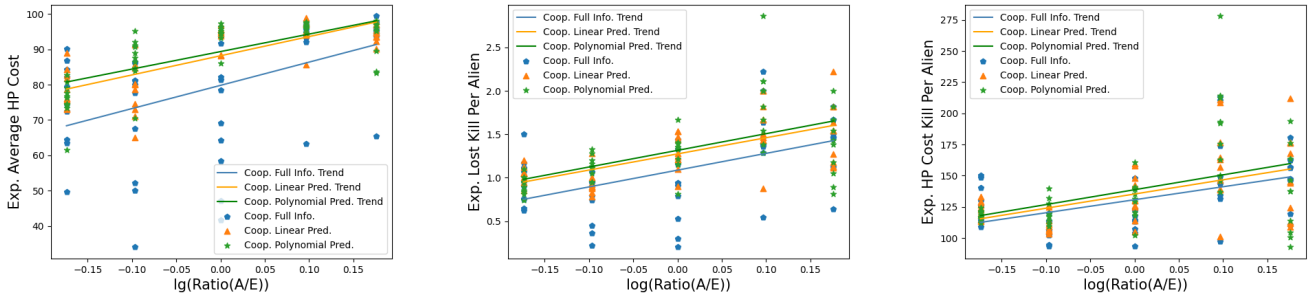
APPENDIX B
MORE DATA ON THE INFORMATION PREDICTION EXPERIMENTS IN ENVIRONMENT WITHOUT OBSTACLES

Fig. 17 shows the performance results for this scenario. From an agent's perspective, the *Linear Regression* model has more accuracy than the *Polynomial Regression* model, comparing with the result trend of *Complete Information* (ground truth). From a system perspective, the win rate and the mean energy/HP cost with different predictive models show similar results. The results in Table. V support this observation.

APPENDIX C
OBSTACLE AVOIDANCE ALGORITHM

When explorers perceive the obstacles (which are unintentional adversaries), they utilize limited information by sharing the perceiving information among agents for cooperative collision avoidance. In the current implementation, robots can $R5$ detect the obstacles using their onboard proximity sensors.

An example scenario with nine robots forming an ally agent team and two mountains representing obstacles on the way to a treasure is depicted in Fig. 18. To avoid a collision, the robot team needs to switch the moving direction k based on the tangent's direction of the nearest collision point c. There are two directions n and m. According to the current status, $R5$ will select the direction n, which has potentially more non-collision robots.



(a) Explorer Average HP Cost (b) Kill Per Alien - Explorer Lost (c) Kill Per Alien - HP Cost

Fig. 17: Performance results of information prediction experiments with different predictive models without obstacles in the environment.

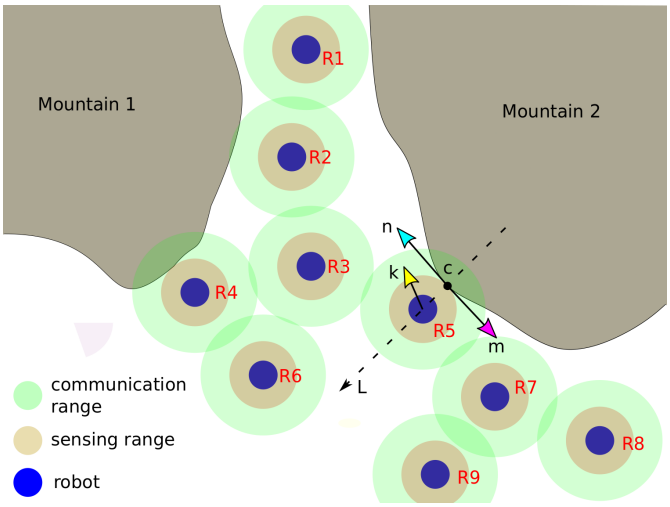


Fig. 18: Illustration of "Adapt The Edge" algorithm for tackling (unintentional) obstacles.

Specifically, L is a straight line passing through the tangent point c and perpendicular to n and m. There are four robots $R1$, $R2$, $R3$, and $R6$ without collision in the direction n comparing with three non-collision robots $R7$, $R8$, and $R9$ in the direction m in current situation. So $R5$ will move Δd in direction n, then adjust the direction to the goal point moving forward until there is no unintentional adversaries in its route through iterating the process. Therefore, the ally robot team adapts the edge of their formation shape dynamically to avoid obstacles. Alg. 2 presents the decision process of the proposed obstacle avoidance strategy.

Appendix D
Utility Values Computation

We implement the agent needs hierarchy through the following levels and parameters:

- The teaming needs are represented using the team *winning* expectation, which is used in the Level 1 payoff matrix of the GUT (Table I) and implemented using the calculations in Sec. D-1 ;
- The basic needs are represented using the (battery) *Energy* level metric, which is used in the Level 2 payoff matrix of the GUT (Table II) and implemented using the calculations in Sec. D-2 ;
- The safety needs are represented using the Health Points (*HP*) status metric, which is used in the Level 3 payoff matrix of the GUT (Table III) and implemented using the calculations in Sec. D-3.

The capability needs, and the learning needs as per Sec. III-B are not implemented in this work.

Below, we list the terms and notations to be used in the below utility equations.

- AT and BT represent the action space of group A and B correspondingly;
- n and m represent the number of Explorers and Aliens respectively;
- d represents the group average distance between two opponents;
- v represents the agent's velocity;
- i and j represent the times of attacks and being attacked;
- w represents Explorers' communication times;
- f and q represent the unit attacking energy cost of both sides agents respectively;
- t_{ev} and t_{mv} represent average attacking ability levels of both sides respectively;
- r_{ev} and r_{mv} represent average defending ability levels of both sides respectively;
- t_e and t_m represent specific agent's attacking ability levels of both sides respectively;
- r_e and r_m represent specific agent's defending ability levels of both sides respectively;
- ϕ_e and ϕ_m represent individual agent's size;
- k represents the number of Explorers' attacking simultaneously;
- g represents the number of Aliens' attacking simultaneously;
- a, b, c, ρ, $\gamma_{e/m}$ and $\delta_{e/m}$ represent corresponding coefficient;
- e_e and e_m represent the current energy level of Explorer and Alien;
- h represents the current HP level of agent;
- p represents the probability values of a specific variable.

1) Winning Utility Expectation: We consider the *Winning Utility* following *Bernoulli Distribution* to represent individual high-level expected utility (teaming & cooperation needs) in the first level (Eq. (33)). The values of the coefficient $a_{1...5}$ are based on different situations in the game.

$$W(t_{ev}, t_{mv}, r_{ev}, r_{mv}, n, m) = (a_1 \frac{a_2 t_{ev} + a_3 r_{ev}}{a_4 t_{mv} + a_5 r_{mv}})^{\frac{m}{n}}; \quad (33)$$

2) Energy Utility Expectation: The second level's utility can be described as the relative *Expected Energy Cost* Eq. (34), which consists of three parts of energy costs: $moving$, $attacking$, and $communication$. And we assume that the energy of $moving$, $attacking$, and $communication$ are proportional to the distance between agents Eq. (35), times of attacks and being attacked Eq. (36), and the communication times Eq. (37), respectively.

$$\begin{aligned} &E(d,v,f,q,n,m,\phi_e,\phi_m) = b_0 + \\ &b_1 \int_{-\infty}^{+\infty} (n-m) e_d(x) p_d(x,d) \mathrm{d}x \\ &+ b_2 (\sum_{i=1}^{+\infty} n e_{a_e}(i,f) p_{a_m}(j, m\phi_m) - \\ &\sum_{j=1}^{+\infty} m e_{a_m}(j,q) p_{a_e}(i, n\phi_e)) \\ &+ b_3 \sum_{w=1}^{+\infty} n e_c(w) p_c(w, \frac{d}{v}); \end{aligned} \tag{34}$$

$$e_d(x) = b_{11} x; \tag{35}$$

$$e_a(x,y) = b_{12} xy; \tag{36}$$

$$e_c(x) = b_{13} x \tag{37}$$

In the attack-defend process, we also assume that individual action distance (Eq. (38)), times of attacks (Eq. (39)) and being attacked (Eq. (40)), and the communication times (Eq. (41)) follow *Normal Distribution*, *Poisson Distribution* correspondingly.

$$p_d(x,d) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-d)^2}{2}}; \tag{38}$$

$$p_{a_e}(x, \lambda_e) = \frac{e^{-\lambda_e} \lambda_e^x}{x!}; \tag{39}$$

$$p_{a_m}(y, \lambda_m) = \frac{e^{-\lambda_m} \lambda_m^y}{y!}; \tag{40}$$

$$p_{a_m}(x, \frac{d}{v}) = \frac{e^{-\frac{d}{v}} (\frac{d}{v})^x}{x!}; \tag{41}$$

Then, we can simplify the Eq. (34) as Eq. (42).

$$\begin{aligned} &E(d,v,f,q,n,m,\phi_e,\phi_m) = \\ &b_0 + b_1 b_{11} (n-m) d + \\ &b_2 b_{12} nm (f\phi_m - q\phi_e) + \\ &b_3 b_{13} n \frac{d}{v}; \end{aligned} \tag{42}$$

3) HP Utility Expectation: At the lowest level, we use the expected HP cost to describe the expected utility Eq. (43), which is the relative HP cost between explorer and alien. And we assume that the unit HP cost is associated with times of being attacked and individual unit attacking and resistant level Eq. (44). Also, individual unit attacking Eq. (45) and resistant Eq. (46)) level is proportional to the agent's current energy level, correspondingly.

$$\begin{aligned} &H(k,t_e,t_m,r_e,r_m,g,\phi_e,\phi_m) = \\ &c_0 + c_1 (\sum_{i=1}^{+\infty} k h(t_e, r_e, i) p_{h_m}(i, \phi_m) - \\ &\sum_{j=1}^{+\infty} g h(t_m, r_m, j) p_{h_e}(j, \phi_e)); \end{aligned} \tag{43}$$

$$h(x,y,z) = \rho z (x+y) \tag{44}$$

$$t_{e/m}(e_{e/m}) = \gamma_{e/m} e_{e/m}; \tag{45}$$

$$r_{e/m}(e_{e/m}) = \delta_{e/m} e_{e/m}; \tag{46}$$

Here, we assume that $p_{h_e}(i, \phi_e)$ and $p_{h_m}(j, \phi_m)$ also follow *Poisson Distribution*. Then, through simplifying the Eq. (43), we finally get Eq. (47).

$$\begin{aligned} &H(k,t_e,t_m,r_e,r_m,g,\phi_e,\phi_m) = \\ &c_0 + c_1 \rho [k\phi_m e_e (\gamma_e + \delta_e) - \\ &g\phi_e e_m (\gamma_m + \delta_m)]; \end{aligned} \tag{47}$$