Doctrine Informed Maneuver in a Large Starcraft2 Environment

Anjon Basak[b], Erin Zaroukian[a], Mark Mittrick[a], John Richardson[a], James Zachary Hare[a], Rolando Fernandez[a], Nick Waytowich[a], and Derrik Asher[a]

[a]DEVCOM Army Research Laboratory, Adelphi, MD, USA
[b]DEVCOM ARL Research Associateship Program, Austin, Texas, USA

ABSTRACT

The performance of decision-making algorithmic approaches depends on a vast number of factors, including hyperparameters, which make some solutions difficult to find. In our previous work (MARDOC paradigm),[1] we have found that even in a simple environment[2] (i.e., a small map with few obstacles), merely changing the initial conditions or doctrinal guided policy (MARDOC) showed a significant impact on the converged behavior. Further, we found that not all policies were useful or desirable for military applications. In this paper, we focus on a complex environment (i.e., a larger map with a greater number of heterogeneous assets and a stronger adversarial force) to analyze the impact of different doctrinal control parameters on the performance and behavior of fixed doctrinal policies. Especially we prioritize the Red force assets for targeted maneuvers and attacks. We hypothesize that asset type and their corresponding coordination will have a significant impact on the performance of the Blue force. Our preliminary experiments in this complex environment showed that the performance varies tremendously depending on asset capability and coordination between teams.

1. INTRODUCTION

Military domains are inherently distributed across large heterogeneous spatial areas[3] (e.g., city) that make it difficult for the commander (e.g., a battalion commander) to make maneuvering decisions. Before making any decision the commander has to gather necessary intelligence (e.g., enemy position, capability, strategy). Even when considering a strategic maneuver the commander has to consider whether there is enough time and space to make the maneuver. Since maneuvering a battalion requires adequate space and time, it is not easy to change the strategy abruptly. Before and after gathering intelligence, a commander utilizes Army doctrine that facilitates a good starting point for making decisions. Army doctrine is time-tested knowledge that has been implemented and executed in a real-world environment. It acts not as a fixed rule but as a flexible starting point for making decisions under multiple complex variables (e.g., terrain, enemy positions, capabilities, etc.).

Frequently, a commander also utilizes different simulation environments and different learning algorithms to inform his decision-making process. However, a simulation environment is not a perfect representation of the real world, and algorithms (e.g., RL algorithms) struggle to perform efficient exploration due to the large state space. On the contrary, if a simulation environment is small and simple compared to the real-world environment, the transition of a simulated and evaluated strategy will not always reflect well in a real environment. Also, we have seen[1] that using learning algorithms does not always converge to a maneuvering strategy that can be utilized in the military domain. So, the problem of larger state space becomes not just an MDMP problem but also a research problem. Our experiments show that it is necessary to incorporate Army doctrine into the decision-making algorithms so that the strategy can give useful guidance to a commander.

In this work, we consider a large Starcraft2 map (256x256) called TigerClaw that is 8 times bigger than usual SMAC[4] maps to evaluate different Army doctrinal maneuvers gathered from experts in the military domain: how the prioritization and suppression of different enemy assets and team coordination impact performance. Our initial results show that by prioritizing Red force assets the Blue force performance increases drastically especially considering casualties. These results also illustrate that it is important to incorporate Army doctrine to make the decision-making algorithms more capable and useful.

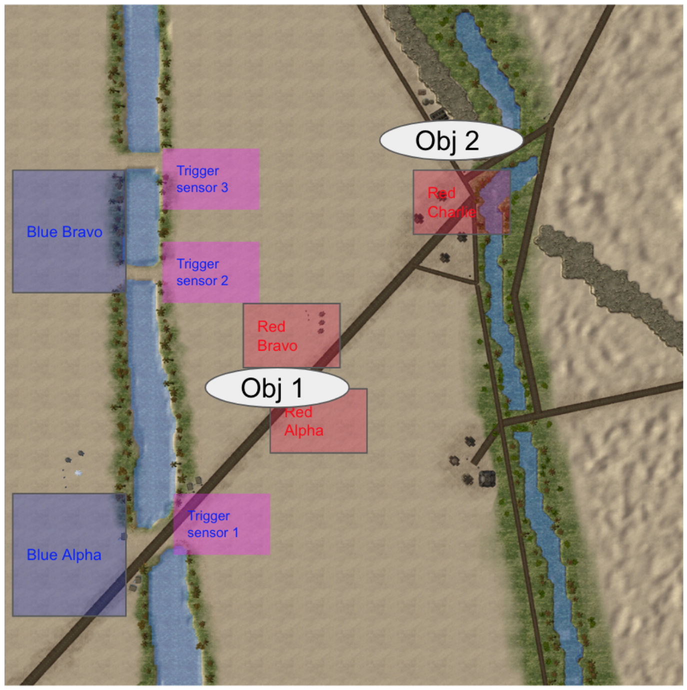


Figure 1: Simplified TigerClaw map

2. METHODS

We utilized a custom Starcraft2 map, called *TigerClaw*, which was derived from a scenario developed by the Army Maneuver Center of Excellence, for the evaluation of different doctrinal strategies. In this scenario, the Blue force tries to cross the river and take control of different regions from the defending Red force. We evaluate three types of doctrinal expert strategies based on different metrics, e.g., proximity, velocity, firepower, and range: i) Baseline ii) Sequential priority iii) Targeted priority. Each scenario lasts 400 time steps.

2.1 The TigerClaw Scenario

The simplified TigerClaw scenario map, shown in Figure 1, consists of a flat desert-like region with a river located on the west side running from north to south. There are three land-bridges that allow for crossing the river. Initially, the Red forces are deployed on the east side of the river while the Blue forces are deployed on the west side of the river. Each of the three crossings have a trigger sensor that alerts the Red force if the Blue force attempts to cross.

The Blue force and Red force both have 16 assets with different capabilities and are divided into teams, summarized in Table 1. The Blue force is divided into *Blue Alpha* and *Blue Bravo*. The Blue Alpha team consists of 3 tanks and 4 helicopters located on the southwest side of the crossings. The Blue Bravo team consists of 3 tanks, 3 infantry, 2 mechanical infantry, and 1 helicopter located on the northwest side of the crossings.

The Red force is heavy on aviation and artillery and is divided into three teams, *Red Alpha*, *Red Bravo*, and *Red Charlie*. Red Alpha consists of 1 artillery siege mode, 1 artillery tank mode, 2 helicopters, and 2 tanks. Red Bravo consists of 1 artillery siege mode, 1 artillery tank mode, 1 helicopter, and 2 tanks. Red Charlie consists of 1 artillery tank mode, 2 helicopters, and 2 mechanical infantry. The Red Alpha and Bravo teams defend the bridges on the east side of the *Obj 1* region. The Red Charlie team located in the *Obj 2* region acts as reinforcement in case breaching starts to happen. In this scenario, if any of the Blue force assets cross the bridges and trigger the sensors, then the Red force attacks all of the Blue force assets (both Blue teams). The strategy followed by the Red force is controlled by the default Starcraft2 heuristic algorithm for attacks.

Blue force:	**Blue Alpha**	**Blue Bravo**	
	3 tanks	3 tanks	
	4 helicopters	3 infantry	
		2 mechanical infantry	
		1 helicopter	
Red force:	**Red Alpha**	**Red Bravo**	**Red Charlie**
	1 artillery siege mode	1 artillery siege mode	1 artillery tank mode
	1 artillery tank mode	1 artillery tank mode	2 helicopters
	2 helicopters	1 helicopter	2 mechanical infantry
	2 tanks	2 tanks	

Table 1: Summary of Blue force and Red force assets.

Blue Alpha	Blue Bravo		Red Alpha	Red Bravo	Red Charlie
4 helicopters	1 helicopter	→ 1)	2 helicopters	1 helicopter	2 helicopters
3 tanks	3 tanks	2)	1 artillery tank mode	1 artillery tank mode	1 artillery tank mode
	3 infantry	3)	1 artillery siege mode	1 artillery siege mode	
	2 mechanical infantry	4)	2 tanks	2 tanks	
		5)			2 mechanical infantry

Table 2: S_heli strategy schematized to show priority of Red force assets.

The purpose of the Blue force is to cross the river and take over the *Obj 1* and the *Obj 2* regions sequentially and destroy the Red force assets in the way. The challenge for the Blue force is that they have only 16 assets against 16 Red force assets. According to doctrine, a 3:1 combat-power ratio (overmatch) is required before engaging the enemy. For our experiments, we have kept the number of assets equal and the ratio of aviation assets 1:1 to illustrate the performance of different doctrinal strategies.

2.2 Doctrinal Approach

Now we present different doctrinal expert approaches that we used as the Blue force strategy against the Red force. The strategies are based on different metrics and locating and prioritizing Red force and then suppressing their firepower so that the rest of the Blue force can maneuver to achieve their objective. This ordering is particularly effective if the Red force has strong aviation and long-distance artillery assets.

2.2.1 Baseline

The *Baseline* approach prioritizes the Red force based on the euclidean distance metric. Each Blue force member computes the closest Red force asset for attack if it is alive.

2.2.2 Sequential Priority

In the *Sequential Priority* strategies, the Blue force locates and then prioritizes the assets in terms of velocity, damage capability, and maximum range.

S_heli: The Red force assets are prioritized in the following order where 1 is the highest priority: 1) helicopter, 2) artillery tank mode, 3) artillery siege mode, 4) tank, 5) mechanical infantry. See Table 2.

S_art: In this strategy the artilleries are given the most priority in the following order: 1) artillery tank mode, 2) artillery siege mode, 3) helicopter, 4) tank, 5) mechanical infantry. See Table 3.

After prioritizing, each of the Blue force team members computes the alive Red force team member according to the priority order and initiates attacks. If the target is not alive then the next target is chosen from the priority queue.

Blue Alpha	Blue Bravo		Red Alpha	Red Bravo	Red Charlie
4 helicopters	1 helicopter	→ 1)	1 artillery tank mode	1 artillery tank mode	1 artillery tank mode
3 tanks	3 tanks	2)	1 artillery siege mode	1 artillery siege mode	
	3 infantry	3)	2 helicopters	1 helicopter	2 helicopters
	2 mechanical infantry	4)	2 tanks	2 tanks	
		5)			2 mechanical infantry

Table 3: S_art strategy schematized to show priority of Red force assets.

Blue Alpha	Blue Bravo		Red Alpha	Red Bravo	Red Charlie
4 helicopters	1 helicopter	→	2 helicopters	1 helicopter	2 helicopters
3 tanks	3 tanks	→ 1)	1 artillery tank mode	1 artillery tank mode	1 artillery tank mode
	3 infantry	2)	1 artillery siege mode	1 artillery siege mode	
	2 mechanical infantry	3)	2 tanks	2 tanks	
		4)			2 mechanical infantry

Table 4: T_heli strategy schematized to show helicopter/the rest and priority of Red force assets.

2.2.3 Targeted Priority

In these strategies, the Blue force team members' actions not only depend on the Red force assets' capability but also their own capability. We divide the Blue force team members into two groups: a) all the helicopters and b) the rest (tank, infantry, and mechanical infantry). This team division does not imply that all the helicopters are located in one place and the others are located in another place. The geographical team division is described in the Maps section. However, the helicopters and the rest of the Blue force assets follow different strategies where the two strategies are primarily differentiated based on their prioritizing of the Red force assets and how the Blue force team members divide their responsibilities between their two teams.

T_heli: In this strategy all the Blue force helicopters target all the Red force helicopters, and after their job is finished they return back to their base. All the other Blue force assets target the rest of the Red force assets: all the Blue force infantry, tank, and mechanical infantry attack all the Red force artilleries, tanks, and mechanical infantries according to the priority queue of **S_heli** strategy. See Table 4.

T_art: All the Blue force helicopters target all the Red force artilleries, and after their job is finished they return back to their base. All the other Blue force assets target the rest of the Red force assets: all the Blue force infantry, tank, and mechanical infantry attack all the Red force helicopters, tanks, and mechanical infantries according to the priority queue of **S_art** strategy. See Table 5.

T_heli_art: This is the same as the *T_heli* strategy however the Blue force helicopters targets all the Red force helicopter and the artilleries in the following priority order: 1) helicopter, 2) artillery tank mode, 3) artillery

Blue Alpha	Blue Bravo		Red Alpha	Red Bravo	Red Charlie
4 helicopters	1 helicopter	→	1 artillery siege mode 1 artillery tank mode	1 artillery siege mode 1 artillery tank mode	1 artillery tank mode
3 tanks	3 tanks	→ 1)	2 helicopters	1 helicopter	2 helicopters
	3 infantry	2)	2 tanks	2 tanks	
	2 mechanical infantry	3)			2 mechanical infantry

Table 5: T_art strategy schematized to show helicopter/the rest and priority of Red force assets.

Blue Alpha	**Blue Bravo**		**Red Alpha**	**Red Bravo**	**Red Charlie**
4 helicopters	1 helicopter	→ 1)	2 helicopters	1 helicopter	2 helicopters
		2)	1 artillery tank mode	1 artillery tank mode	1 artillery tank mode
		3)	1 artillery siege mode	1 artillery siege mode	
3 tanks	3 tanks	→ 1)	2 tanks	2 tanks	
	3 infantry	2)			2 mechanical infantry
	2 mechanical infantry				

Table 6: T_heli_art strategy schematized to show helicopter/the rest and priority of Red force assets.

Blue Alpha	**Blue Bravo**		**Red Alpha**	**Red Bravo**	**Red Charlie**
4 helicopters	1 helicopter	→ 1)	1 artillery tank mode	1 artillery tank mode	1 artillery tank mode
		2)	1 artillery siege mode	1 artillery siege mode	
		3)	2 helicopters	1 helicopter	2 helicopters
3 tanks	3 tanks	→ 1)	2 tanks	2 tanks	
	3 infantry	2)			2 mechanical infantry
	2 mechanical infantry				

Table 7: T_art_heli strategy schematized to show helicopter/the rest and priority of Red force assets.

siege mode. The rest of the Blue force assets (infantry, tank, mechanical infantry) targets the Red force tanks and mechanical infantries in the priority order of **S_heli** strategy. See Table 6.

T_art_heli: This is the same as the *T_art* strategy however the Blue force helicopters targets all the Red force helicopter and the artilleries in the following priority order: 1) artillery tank mode, 2) artillery siege mode, 3)helicopter. The rest of the Blue force assets (infantry, tank, mechanical infantry) targets the Red force helicopter, tanks, and mechanical infantries in the priority order of **S_art** strategy. See Table 7.

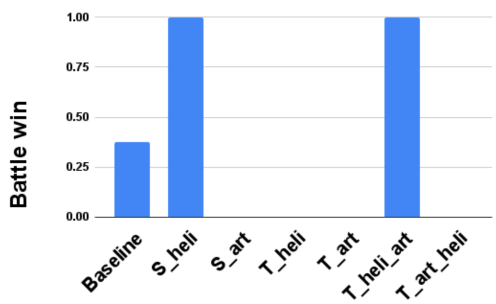
3. RESULTS

In this section, we evaluate the performance of different doctrine-based heuristic strategies in a large simulated environment where learning algorithms fail to learn any useful strategy. Our goal is to show that locating, prioritizing, and suppressing the Red force assets and better coordination among Blue force teams can result in better performance. The result can also give insights into team configuration, responsibilities, and temporal maneuvers which can inform the Course of Action in MDMP.

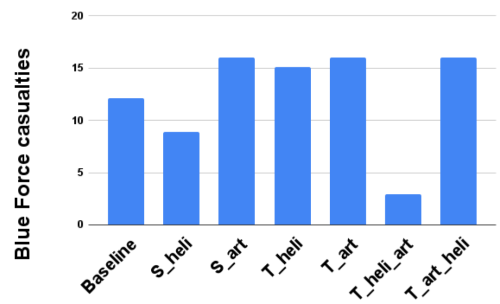
Our results in Figure 2 show that *T_heli_art* and *S_heli* performed the best according to battle win rate and Red force casualties. However, *T_heli_art* was able to perform better in terms of Blue force casualties by maintaining the same battle win rate and Red force casualties of *S_heli*. This is due to *T_heli_art* dividing the responsibilities and better team coordination among the Blue force aviation assets (e.g. helicopters) and ground units (e.g. infantry, tank). In *S_heli* there was no team coordination and every Blue force asset was performing the same prioritized tasks resulting in higher Blue force casualties compared to *T_heli_art*.

By comparing *T_heli_art* with *T_art_heli* we see that it is important to destroy the Red force aviation assets first and then the artillery assets. This is due to aviation assets having better velocity and agility.

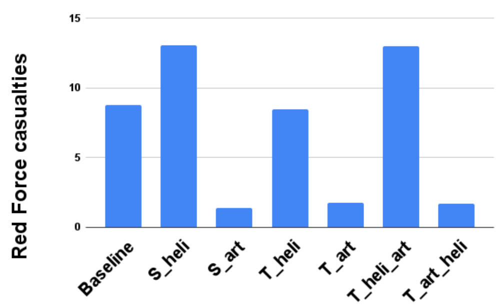
Even though *T_heli*, where the Blue force aviation assets target only Red force aviation assets, did not perform well according to battle win rate. However, by contrasting it with *T_heli_art* we learn: it is important for the Blue force aviation assets to destroy not only the Red force aviation assets but also artillery assets that have a long distance range.



(a) Battle win rate (1=100% of battles won).



(b) Blue force casualties.



(c) Red force casualties.

Figure 2: Comparison between the baseline and doctrine-based approaches showing that *T_heli_art* performing the best achieving maximum win rate, lowest Blue force casualties and maximum Red force casualties.

4. CONCLUSION AND FUTURE WORK

A large state space creates a huge bottleneck for any decision-making algorithm. Due to the inherent characteristics of the environment of the military domain, it becomes a problem not only from the military decision-making perspective but also from the research perspective. Our results show that the simple *Baseline* algorithm does not perform well if we prioritize Red force assets based on just proximity. It is also important to have better team configuration and divide responsibilities between the teams according to their capability. Our results also show that Red force aviation assets should have higher priority than the artillery assets even though it has long range. For our future work, we would like to incorporate this doctrinal knowledge into RL algorithms to improve their efficiency in exploring large state spaces.

Keywords: Reinforcement Learning, MARDOC, Doctrine, StarCraft II Multi-Agent Challenge Environment, Simulation

REFERENCES

[1] Basak, A., Zaroukian, E. G., Corder, K., Fernandez, R., Hsu, C. D., Sharma, P. K., Waytowich, N. R., and Asher, D. E., "Utility of doctrine with multi-agent RL for military engagements," in [*Artificial Intelligence and Machine Learning for Multi-Domain Operations Applications IV*], Pham, T. and Solomon, L., eds., **12113**, 1211323, International Society for Optics and Photonics, SPIE (2022).

[2] Samvelyan, M., Rashid, T., de Witt, C. S., Farquhar, G., Nardelli, N., Rudner, T. G. J., Hung, C.-M., Torr, P. H. S., Foerster, J., and Whiteson, S., "The StarCraft Multi-Agent Challenge," *CoRR* **abs/1902.04043** (2019).

[3] Army, U., "FM-30 - operations." https://armypubs.army.mil/epubs/DR_pubs/DR_a/ARN36290-FM_3-0-000-WEB-2.pdf (Oct 2022). Accessed: 2023-02-18.

[4] Wang, T., "SMAC - starcraft multi-agent challenge." https://github.com/oxwhirl/smac (Oct 2020). Accessed: 2022-02-18.