University of Bristol
Faculty of Engineering
Robotics Research Preparation (AENGM0029)

Research Proposal

Designing Considerate Swarms

Nishant Ramakuru
1977959
MSc Robotics

Supervisor: Namid Shatil
Co-supervisor: Sabine Hauert

Bristol, March 2020

Table of Contents

1 Aims and Objectives **ii**

1.1 Aim ii

1.2 Objectives ii

2 Motivation **iv**

3 Literature Review **v**

4 Risk register **vii**

5 Timeline **viii**

Bibliography **x**

Abstract

Swarms of autonomous agents acting towards a common purpose have been shown to benefit when the individual agents are pre-designed to exhibit elements of cooperation and collaboration. In this project, we will extend this approach to include considerate behaviour; swarms that actively avoid inconveniencing others within the group. To achieve this, we will create swarm algorithms that include an extension to traditional game theory by conditioning each individual action on the preference of other individuals within the group. These algorithms can then be tested in a variety of scenarios, both through simulations and a swarm robotic platform.

Index: Traditional Game theory (TGT), Conditional Game Theory (CGT)

1. Aims and Objectives

1.1 Aim

The aim of the project is showing how the use of game theory can benefit a swarm of intelligent agents and understand how robots will interact and update their strategies to align with the common goal of the swarm.

The system can be modelled at two levels; *Macroscopic*, where the entire swarm is controlled, and *Microscopic*, where each individual agent is controlled [1]. In both the methods, the interaction of the agents is not exclusively taken into consideration. Swarm algorithms rarely include the rule of taking the preferences of other agents of the swarm, or "Considerate behaviour". We define considerate behaviour as the additional rule of conditioning each agent with the actions of other agents.

The proposed method is a mixture of the two modelling scales, where the goal of the swarm will be set at the macroscopic level but an additional rule of conditioning each agent will be carried out at the microscopic level. This is different to the traditional game theory (TGT) notion of the payoff structure, as the payoff structure is designed to be conditional on the players option, so-called conditional game theory (CGT) [2]. To clearly observe the behaviour of the agents, a two agent system will be considered in this project as a starting point.

1.2 Objectives

- **Study and understand TGT**
  The first phase of the project will be to get a thorough understanding of TGT and its components.
- **Build base algorithm for TGT**
  The next step is to understand the mathematical relations and build a base algorithm using Bayesian networks and TGT models. The proposed method will focus on how to incorporate these strategies and models into a swarm algorithm.
- **Simulate TGT algorithm**
  Once the underlying concept and rules of the method are formed, the algorithms will be tested in simulated environments.The proposed method will simulate the TGT algorithm and observe the behaviour of the agents.
- **Incorporate additional rule(considerate behaviour)**
  The next step is to understand the extensions to TGT, known as CGT. Add the additional rule of conditioning each agent with the action preferences of others, using concepts from Bayes theorem. The successful execution of the algorithm will lead to a calculation of the Nash equilibrium for the swarm.

- **Simulate Considerate Behaviour**
  The proposed method will simulate the CGT algorithm and observe the functioning of the agents. The agents in the simulation should exhibit considerate behaviour and should accomplish the goal of the swarm simultaneously.
- **Reinforcement learning (benchmark)**
  An approach to the same task using reinforcement learning will also be carried out as a benchmark to test the performance of the proposed method. A reward system will be incorporated in the reinforcement learning algorithm, where the positive reward will be awarded for exhibiting properties of considerate behaviour and the accomplishment of the task. Negative rewards will be awarded if the agent somehow diminishes the performance of the other agent or does not accomplish the given task.
- **Simulate Reinforcement learning**
  The algorithm will be iterated and trained until convergence. It will be simulated in the same environment as the CGT model to normalize the parameters for comparison.
- **Hardware implementation**
  The final stage of the project will be to enable robots to exhibit the developed strategies in a physical environment, which motivate changes to the algorithm to make it more robust and pragmatic.

2. Motivation

Swarm algorithms usually have two approaches i.e., program agents individually or program the system as a whole. Swarm algorithms rarely incorporate game theory exclusively, as a result there is relatively little research on the intersection of swarm intelligence and game theory.warm algorithms have never incorporated game theory exclusively. The interactions that occur during the functioning of a swarm are dynamic and have not been entirely considered as a governing rule for swarm algorithms. Motivation of this project arises from this relatively unexplored area. The proposed method is an ideal amalgamation of these approaches, where the agents are programmed individually with an additional rule: condition each individual action on the preference of other individuals within the group.

Swarm robotics is often heavily dependent on coordination and not on the performance of a single agent. Performance is the individual efficiency of the agent. It is a good metric with respect to individual assessment but in a team/swarm, co-ordination is much more relevant and important to the system. For example, a group of players who perform very well individually may perform very poorly as a team, whereas a team of players who perform poorly individually may perform very well as a team.

The main idea behind swarm robotics is that, it is a system with large number of rational agents. The friction between agents, or in this case a large number of agent can lead to a drastic drop in performance of the swarm as a whole. Hence the need for cooperation between the agents is highly desirable. This metric of cooperation between agents can be observed, understood and determined by the concepts of game theory. Hence this project will explore the effectiveness of using game theory in swarms.

3. Literature Review

Swarm robotics is a system which consists of large numbers of mostly simple physical robots. The collective behaviour emerging from interactions between robots is referred to as a swarm algorithm. Game theory is the study of mathematical models of strategies based on the interactions of rational agents. The objective of this project is to implement strategies using game theory and an extension to game theory, whereby rational agents increase the performance of the group, or in this case, a swarm. Hence the prerequisites of this project require a thorough understanding of game theory, swarm algorithms and swarm simulation platforms. It also requires a basic understanding of machine learning and artificial intelligence to implement reinforcement learning as a benchmark to asses efficiency of the proposed method.

The following points briefly summarizes the literature.

- **Game theory in robotics:** Coordination vs performance shows us how important game theory can be in swarms [2] . Having intelligent agents with high individual performance with no cooperation can lead to a poorly functioning swarm. On the other hand a poor performance yielding agents with high cooperation can emerge as a good functioning swarm.

  Game theory is already being used in robotics for example, see [3] [4] which give examples of how game theory can be useful in human robot interaction, which is nothing but a two agent swarm. A more complex model, gives a example of how game theory can be used by classifying the swarms as predator and prey [5] shows us that the interpolation of game theory can be carried out in two ways. Either the agents have conflicting goals [5], or the agents work towards a common one [3]. The latter of the two, where the use of game theory is used to engineer considerate behaviour, is desired aim for proposed method.

- **Traditional (TGT):** Two classic examples of the type of scenarios that are widely used in TGT are the "Prisoner's Dilemma" [6] and "Battle Of Sexes" [7], which both refer to two agent systems. This project requires the understanding Nash equilibrium [8], which states how the strategies of the agents should be chosen in order to maximise the output of the total system of agents. This will act as the convergence measure for the algorithm.

- **Conditional Game Theory (CGT):** Conditional game theory, is an extension to traditional game theory. The underlying concept is to relate the actions of agents to the preferences of other agents belonging to the swarm. Mathematical analysis of CGT is outlined in, for example [9] [2].

- **Theoretical Ground for simulations:** Mathematical models to build simulations. Next phase of the project will be to cover CGT and prepare the ground for the theoretical simulations.

- **Simulations:** Simulations are key to this project as it will be the most time consuming and tedious task. Using simulation platforms rather than inventing from scratch will be ideal in this scenario. Platforms like ARGoS will be a good starting point to explore the features and capabilities of softwares [10]. See, for example [11] for reference to comparison of various modelling and simulation methods as per the specifications of the project.

- **Reinforcement learning:** A reinforcement learning technique implemented on a two agent system, similar to the one being proposed will be implemented to make comparison easier. Reinforcement learning in swarms with two agent system such as [12] will act as a guidance to carry out this objective. An actor critic relationship much like the predator-prey [5] relationship is implemented in this paper.

4. Risk register

This section outlines the risks involved in the project and the mitigating steps to be taken in order to reduce them as much as possible.

The following scoring system is used in the risk register table is given below:

- Likelihood: 1 = not very likely; 4 = very likely
- Impact: 1 = not hardly any impact; 4 = major impact
- Risk score = (Likelihood) x (Impact)

Risk	Mitigation	Likelihood	Impact	Score
COVID-19. Leading to limited access to BRL facilities.	Implement algorithms on much simpler hardware equipment purchased online.	4	4	16
Overload due to simulations.	Offload the simulations on cloud computing platforms, such as AWS to run on much faster systems.	3	4	12
Difficulties in hardware implementation	Start implementing each algorithm simultaneously instead of waiting to finish all of them and implement them at once	4	2	8
Longer to generate mathematical models	Getting into literature review prior to the start date and simultaneously start working on learning/creating simulation platforms to avoid delay.	2	3	6
Algorithm does not converge	Hyper tuning parameters and explore all possible configurations.	2	2	4.

5. Timeline

This section outlines the timeline of the project. It has two categories of activities/tasks and milestones. Each work package has several tasks and a milestone/deliverable to determine the completion of the particular work. This will make the easier to monitor and manage the workload and constantly re-evaluate to keep the project on track.

The project is broken down into six work packages. The assigned deliverables to the work packages are shown in the table below:

Work package	Deliverables
1. Understand game theory and its components	1. theoretical base for mathematical models
2. TGT models	2. Simulation results for TGT models
3. CGT models	3. Simulation results for CGT models
4. Reinforcement learning models	4. Simulation results for Reinforcement learning models
5. Hardware implementation	5. Successful implementation of both models
6. Comparison of models	6. Determine best model
7. Write up	7. Dissertation report

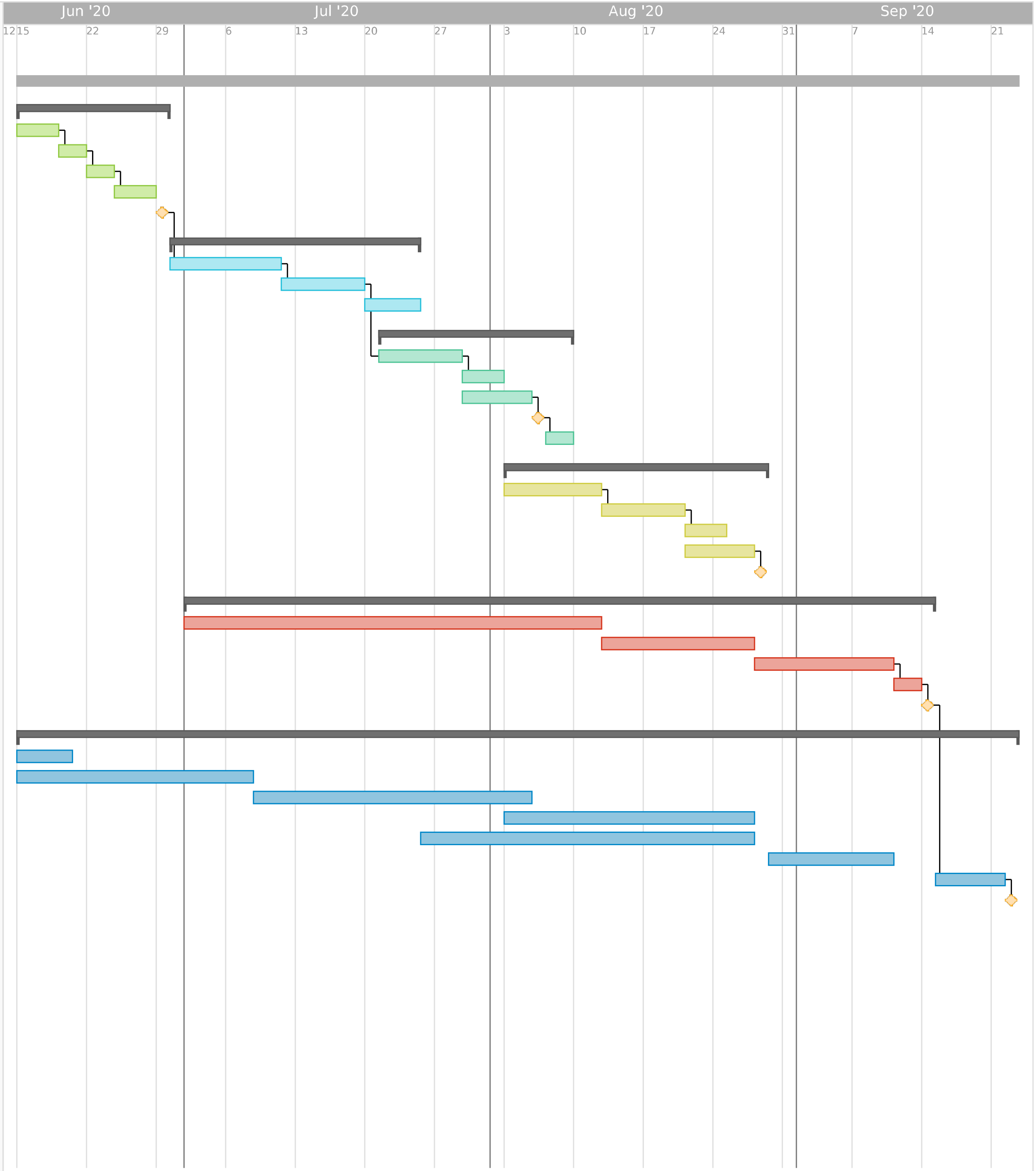
The classification of work packages and deliverables into tasks with respect to time can be seen in the Gantt chart shown below. The timeline is designed such that adequate time is given to each task and the risk of each task is taken into consideration. Overall workflow of the project is the main purpose of this timeline to get a clear picture of the progress. Unrelated tasks have simultaneous implementation, so a delay or a set back in one will not hinder the other. Tasks with simultaneous implementation have been given more time than needed to add flexibility, which makes the workload manageable. The dependent tasks are highlighted with directional arrows for ease of tracking.

The first work-package, although seems less for one of the major tasks of the project i.e, comprehension of the underlying concepts, the necessary work needed to be done is already in progress. The time allotted in the timeline is to glance the concepts before formulating the mathematical models necessary for the next work-package.

The start date of the project is considered as 15th July. The final deliverable of the project is set on the 22nd of September. A slight leeway for each task is considered, due to any unwanted discrepancies, and sufficient surplus time is allotted in each case to make to make sure the deadlines are met.

Dissertation timeline

	start	end
Get a through understanding of g...	**15/06/20**	**29/06/20**
Understand game theory and its co...	15/06	17/06
Understand mathematical equations	18/06	19/06
Conditional Game theory	22/06	23/06
Nash equilibrium	24/06	26/06
Complete mathematical equations	29/06	29/06
Traditional game theory models	**30/06/20**	**23/07/20**
Mathematical models for algorithms	30/06	09/07
Write up algo for game theory	10/07	17/07
Simulate traditional game theory	20/07	23/07
Conditional Game theory models	**21/07/20**	**07/08/20**
Addition of Conditional rule of math...	21/07	28/07
Simulate conditional game theory	29/07	31/07
Observe simulation and tweak para...	29/07	04/08
Nash equilirium achieved	05/08	05/08
Evaluation metrics	06/08	07/08
reinforcement learning models	**03/08/20**	**27/08/20**
Mathematical models for algo	03/08	11/08
Write up algo for reward based syst...	12/08	19/08
Run several iterations for model to l...	20/08	24/08
Tweak parameters	20/08	26/08
Evaluation metrics	27/08	27/08
Hardware implementation & Comp...	**01/07/20**	**14/09/20**
Familiarization with off-the-shelf robo...	01/07	11/08
Test CGT model	12/08	26/08
Test reinforcement model	27/08	09/09
Compare Methods	10/09	11/09
Determine Best method	14/09	14/09
Write up	**15/06/20**	**22/09/20**
Define Outline of project	15/06	18/06
Introduction & Literature review	15/06	07/07
Conditional game theory	08/07	04/08
Reinforecment learning	03/08	26/08
Simulations	24/07	26/08
Hardware Implementation	28/08	09/09
Conclusion	15/09	21/09
Submission	22/09	22/09



Bibliography

[1] Erol Şahin, Sertan Girgin, Levent Bayindir, and Ali Emre Turgut. Swarm robotics. In *Swarm intelligence*, pages 87–100. Springer, 2008. ii

[2] Wynn C Stirling and Villino Volterra. Theory of coordinated agency. *Brigham Young University (mimeo)*, 2016. ii, v, v

[3] Y Li, G Carboni, F Gonzalez, D Campolo, and E Burdet. Differential game theory for versatile physical human–robot interaction. *Nature Machine Intelligence*, 1(1): 36–43, 2019. v, v

[4] Y. Li, K. P. Tee, R. Yan, W. L. Chan, and Y. Wu. A framework of human–robot coordination based on game theory and policy iteration. *IEEE Transactions on Robotics*, 32(6):1408–1418, 2016. v

[5] Haibin Duan, Pei Li, and Yaxiang Yu. A predator-prey particle swarm optimization approach to multiple ucav air combat modeled by dynamic game theory. *IEEE/CAA Journal of Automatica Sinica*, 2(1):11–18, 2015. v, v, vi

[6] William Poundstone. *Prisoner's Dilemma/John von Neumann, Game Theory and the Puzzle of the Bomb.* Anchor, 1993. v

[7] Martin J Osborne and Ariel Rubinstein. *A course in game theory.* MIT press, 1994. v

[8] John F. Nash. Equilibrium points in n-person games. *Proceedings of the National Academy of Sciences*, 36(1):48–49, 1950. ISSN 0027-8424. doi: 10.1073/pnas.36.1.48. URL `https://www.pnas.org/content/36/1/48`. v

[9] Wynn C Stirling. Conditional coordination games on cyclic social influence networks. *IEEE Transactions on Computational Social Systems*, 6(2):250–267, 2019. v

[10] C. Pinciroli, V. Trianni, R. O'Grady, G. Pini, A. Brutschy, M. Brambilla, N. Mathews, E. Ferrante, G. Di Caro, F. Ducatelle, T. Stirling, Á. Gutiérrez, L. M. Gambardella, and M. Dorigo. Argos: A modular, multi-engine simulator for heterogeneous swarm robotics. In *2011 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 5027–5034, 2011. vi

[11] Ying Tan and Zhong-yang Zheng. Research advance in swarm robotics. *Defence Technology*, 9(1):18–39, 2013. vi

[12] J. F. Peters, C. Henry, and S. Ramanna. Reinforcement learning in swarms that learn. In *IEEE/WIC/ACM International Conference on Intelligent Agent Technology*, pages 400–406, 2005. vi