

Final Report

Self-Organising Multi-Agent Systems

Department of Electrical and Electronic Engineering
Imperial College London

SOMAS Class 2021-2022

January 2022

Lecturer

Prof. Jeremy Pitt

Abstract

Nulla malesuada porttitor diam. Donec felis erat, congue non, volutpat at, tincidunt tristique, libero. Vivamus viverra fermentum felis. Donec nonummy pellentesque ante. Phasellus adipiscing semper elit. Proin fermentum massa ac quam. Sed diam turpis, molestie vitae, placerat a, molestie nec, leo. Maecenas lacinia. Nam ipsum ligula, eleifend at, accumsan nec, suscipit a, ipsum. Morbi blandit ligula feugiat magna. Nunc eleifend consequat lorem. Sed lacinia nulla vitae enim. Pellentesque tincidunt purus vel magna. Integer non enim. Praesent euismod nunc eu purus. Donec bibendum quam in tellus. Nullam cursus pulvinar lectus. Donec et mi. Nam vulputate metus eu enim. Vestibulum pellentesque felis eu massa.

Contents

1	Introduction	1
1.1	Problem Statement	1
1.2	Our Goals	1
2	Organisation Structure	2
3	Simulation Structure	3
3.1	Simulation Environment	3
3.2	Simulation Flow	3
3.3	Message Passing	3
3.4	Health Decay System	3
4	Visualisation	4
5	Team 2 Agent Design	5
6	Team 3 Agent Design	6
6.1	The Agent	6
6.1.1	Action Process	7
6.1.2	Agent Knowledge	8
6.1.3	Agent Generation	8
6.2	Agent Communication	9
6.2.1	The Agent's Social Network	9
6.2.2	Message Sending	10
6.2.3	Message Reception	10
6.2.4	Reaction to Received Messages	10
6.2.5	Proposing a Treaty	11
6.2.6	Handling a Proposed Treaty	12
6.3	Food Intake	12

7	Team 4 Agent Design	14
8	Team 5 Agent Design	15
9	Team 6 Agent Design	16
10	Team 7 Agent Design	17
11	Experiments	18
12	Results	19
13	Discussion	20
14	Conclusion	21
15	Future Work	22
A	Appendix	23
	Bibliography	24

Chapter 1

Introduction

1.1 Problem Statement

Section on the tower and how we have interpreted the spec. Should include a basic spec for features that we wanted to achieve

1.2 Our Goals

Getting a simulation environment working

Chapter 2

Organisation Structure

Could introduce each of the teams and a one liner about their agent strategy. This will obviously be expanded upon in the agent sections. Should contain an org chart and describe important parts of it. Justify certain organisational choices. Maybe consider what could have been done better?

Chapter 3

Simulation Structure

3.1 Simulation Environment

Goes over the backend design

3.2 Simulation Flow

The order in which events happen Need a nice diagram for this

3.3 Message Passing

Need a nice diagram for this

3.4 Health Decay System

Need a nice diagram for this

Chapter 4

Visualisation

Important things about front end.

Chapter 5

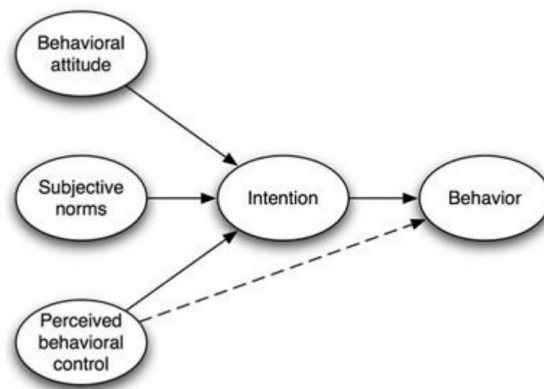
Team 2 Agent Design

Chapter 6

Team 3 Agent Design

6.1 The Agent

Team 3's agents are based on the Theory of Reasoned Action (TRA). The agents are designed to have a shifting personality dependent on the three variables that this theory links to the behaviours of an agent. These three variables, together with a social network, define their personality as a whole, Ocaya [1985] and influence the actions they will take.



The Theory of Reasoned Action was developed by Martin Fishbein in 1967. The TRA states that the behavioural intentions of individuals can be attributed to three interdependent factors: Perceived behaviour control, the subjective norm, and the attitude of the individual, as you can see in the figure above, Norazlan Hasbullah [2014]. Although it was recognised from the start, it must be noted that the TRA is unable to accurately and consistently predict non-volitional behaviours of the individual. Therefore, TRA can only be applied to those actions in which a reason can be found, actions that occur as a reaction to others and that the individual has control over, Fishbein [2008]. This is applicable to our scenario, in which the agent must make a decision on how much food to eat after deliberating on their social network's messages.

The application of TRA, in this case, is ideal given that reversing the methodology enables us to define an individual through the 3 core personality traits. The theory is a valid methodology to predict future behaviour by observing occurring actions and finding those components that caused them. Those components of behaviour can also be extrapolated to define an individual's behaviour and used to define it from the start. Therefore building the main characteristics of the personality of the individual, although lacking the uncontrollable or instinctive reactions.

Following the Theory of Reasoned Action, we have defined the agents' emotion variables to be:

1. Stubbornness: it addresses perceived behaviour control, which is the influence that external agents have on the decisions that are to be taken. It is applied by the agent to ignore or read incoming messages and succumb or not to group pressure.
2. Morality: it addresses the subjective norm, which is one's idea of what is socially perceived to be right or wrong.
3. Mood: it defines the attitude towards the action to take.

These variables are modified as agents interact, which lead our agent to make different decisions during its time in the tower. All possible changes to these variables are done through ranges with the exact change each time being randomised, in other words inside a specified range any of the variables will change a different amount as a reaction for the same action. This course of action is based on the substantive uniqueness of each behaviour as defended by Martin Fishbein, Fishbein [2008]. This characteristic of randomness in the agents' behaviour fulfilled the role of covering the fact that TRA does not take into account non-volitional behaviours, making the agent act as much as possible in a human-like behaviour.

The next step after identifying our agent strategy was to think about the scenario and how the problem should be solved. Our first approach used classical game theory, which states that under normal circumstances, an unbiased agent would always prioritise themselves and their needs over anything else. Hence, this dictates that each agent should not act in a moral way at all. In our scenario, this leads to agents taking selfish amounts of food, causing agents below to have less food and potentially die. However, we found research demonstrating how humans in a social environment tend to make decisions benefiting the collective, Batson et al. [1995]. This contradicted traditional game theory views and made this approach unreasonable when trying to recreate a human type behaviour, following the two most common explanations of why humans violate classic game theory: enlightened self-interest and group identity.

6.1.1 Action Process

The Agent follows the following steps during their day:

1. The agent starts by updating their defining variables at the beginning of each day based on their HP. This acts as the agents checking in on themselves and changing their attitude(mood) and morality depending on their condition(HP).
2. If the agent was just 'born' it will estimate an initial value for the average food consumed.
3. When the tower reshuffles the agents (which each agent knows by checking what floor they are on), the agent's mood changes depending on the relative position of the new floor to the other floor.
4. The agent attempts to eat every day. The quantity depends on treaties, the messages they have received and whether other agents have requested them to leave or eat a certain amount of food or a calculation based on the agent's current values. More information on this can be found in Section 6.3.
5. Finally, the agent will attempt to communicate with other agents and create a social network via message sending and treaty proposition. More information on this is can be found in Section 6.2.

6.1.2 Agent Knowledge

Agent 3 will be able to remember facts in order to make decisions. The information that is stored in their memory is:

1. The floors the agent has been in before. This information will affect the agent's mood after being reshuffled, as they would remember where they have been and know if the new floor is a good or bad one.
2. The last recorded HP value. Used to see changes in the agent's health.
3. The agent's friendship network. They are aware of the people they have met during their time in the tower. The agent's friends are identified and are considered more or less close to the agent with an affection parameter that ranges from 0 to 1, 0 being complete dislike for the other agent.
4. The amount of food last eaten.
5. The amount of food last seen on the platform.
6. A moving average of food consumed.
7. The age of the agent.
8. The treaty we have proposed until it is accepted or rejected.
9. The estimation of the reshuffle.
10. The HP our neighbour above last told us they had.
11. The HP our neighbour below last told us they had.

6.1.3 Agent Generation

The defining variables of our agents are randomised when they are generated in order to avoid preconditioned agents and mimic different personalities. All three variables are defined to be in the range 0 to 100. Morality and mood are randomly allocated to a value in the range, but stubbornness is limited to a number between 0 and 75. As explained in our agent parameters, stubbornness represents the effect of external influences on an agent. A stubbornness of 100 would mean this agent is completely unwilling to listen and would ignore all incoming messages, effectively making them an agent that disturbs the tower. In order for this to not be the case, we added the limitation to the value of stubbornness.

As Shao et al. have shown in "Beyond Moral Reasoning: A Review of Moral Identity Research and Its Implications for Business Ethics", Shao et al. [2008], one's moral action is most heavily influenced by their moral personality, as opposed to their moral reasoning, which empirically seems to show only a minor influence on the moral behaviour of an agent, Blasi [1983]. Hence, we used this variance and unpredictability to initialise our agents randomly on the moral scale. This insight also allowed us to witness a purely emergent strategy and nature as opposed to a pre-programmed inherent morality.

While the initial randomness allowed for interesting strategies to occur, we also wanted our agents to adapt their morality on the basis of their community and associations, as argued by Shao. Making it possible by having all actions an agent does affect its emotion variables.

6.2 Agent Communication

Communication between agents is fundamental for the possibility of collaboration and self-organisation between agents. Crowd behaviour is considered to be generated by individuals, context-dependent and dynamic, as defended by the modern foundation of crowd behaviour, Wijermans [2011]. In this project, communication between agents is based on two concepts, treaties and direct messages, which are equally exploited by our agent and combined with the conceptualization of friendship levels.

6.2.1 The Agent’s Social Network

The agent forms relationships with others surrounding them as they communicate with each other. The agent remembers everyone they have met and assigns a level of friendship to them depending on their interactions. By doing this, the agent forms a social network that will affect their decisions, something that the Theory of Reasoned Action explores in the perceived behaviour control aspect of the theory. It is shown that all three main variables are linked together and for this reason, the agent will have a better or worse opinion of someone they have just met depending on their morality at that point.

Relationships are of high esteem for the agent as it follows the Aristotelian concept of friendship which defends friendship as “the finest way we exercise practical reasoning”. Quoting Aristotle: “Friendship seems to hold cities together, and lawgivers seem to care for it more than for justice ... and when men are friends, they have no need for justice”. Friendship is considered the best form of human moral action by The Nicomachean Ethics, Sokolowski [2002], and is fundamental in the decision-making process of the agent and its response to communication started by other agents.

McCullough et al. investigate thoroughly the nature of prosocial action and gratitude’s place in transactions, McCullough et al. [2008]. While this is difficult to realise fully in the code, we aimed to use the definitional guidelines set out, to relate our agent’s action and emotion to some of the more nuanced but socially beneficial emotions. Our friendship function allows us to assign a level of gratitude to our agent in response to the prosocial action of the other agents. In this way, we could more effectively alert the agent to the prosocial nature of another’s actions. By utilising the friendship function to, in turn, behave more prosocially to those who have done so to us, we embedded a system of gratitude in our agents’ actions.

We graduated our response by negotiating the selflessness of the action of the foreign agent. That is to say, in our game, the more severe the sacrifice on the behalf of the foreign actor, the more grateful we became. This is based on McCullough’s definition that one experiences gratitude when “they have benefited from someone’s costly, intentional, voluntary effort on their behalf”.

Emotions allow us to respond to reality in a self-protective and self-enhancing way, Ekman [1992], it was this response we wanted our agent to have so that the emotions would govern their actions, by either leading them on a self-protective path, if circumstances such as selfish agents prevailed, or a self-enhancing manner, as cooperation would beget further cooperation through our system of gratitude.

This is the basis of our friendship metric and implies our agent will react more positively to a friend and will increase the friendship of those with whom it has positive interactions.

6.2.2 Message Sending

Message sending is determined by a probabilistic decision, with each message having a 0.2 probability of being sent. The direction the messages are sent in is determined by the moral state of the agent. When the agent's morality is low, they will tend towards self-preservation and send messages upwards requesting information or help. On the other hand, when morality is high, the agent will look to create a stable society as stated by their willingness to look ahead and communicate with those below him to offer help and information, aiming to fortify their social network.

The design of message passing has various noteworthy characteristics that act to aid the agent and occur if the conditions make it necessary:

1. The agent prioritises itself when its HP is critical, opting for treaties and messages that involve other agents leaving food in the platform and utilising this mechanic as a failsafe when in danger.
2. The agents take into account the neighbours health when sending messages proposing how much food to leave on the platform. This makes the message more likely to be accepted.

6.2.3 Message Reception

When messages are involved, stubbornness is used to decide to ignore or read a message, with our reaction or answer also being dependent on the stubbornness of the agent.

The agent will follow the next process when a message is received:

1. The agent will decide to ignore or read the message, with a probability of ignoring the message equal to the stubbornness level.
2. If the agent decides to read the message it is analysed and matched with the type of response or responses acceptable for the message.
3. The agent will note the agent that sent the message and consider the friendship level it shares with the sender. Friendship together with the morality level and the content of the message itself will decide the exact content of the response.
4. The content of the message and its sender will affect the agents' emotional variables and the friendship level with the sender. This change affects the future reaction to communication done by other agents and the decisions taken by the agent from that point onwards.

To ensure a minimum level of communication all messages read will be answered by the agent.

6.2.4 Reaction to Received Messages

The changes are dependent on four criteria to determine initial agent characteristic manipulation. These four variables are the variables present upon the receipt of the message and somewhat representative of the type 1 brain system discussed in Kahneman and Tversky [2000], Frankish [2010].

The initial emotional response is based on the variables of the immediate past, the hp moving average on the last five ticks, opposed to the overall hp average, or welfare average. This is to

represent the subconscious layer of emotional response, as seen in, Kamble [2021], whereby our response to the situation is more irrational and illogical than the conscious mind perceives.

Summing the net total of all the initial responses to zero, allows us to explore the oscillatory nature of emotion while avoiding the possible pitfall of any dominant emotions.

HP dec last tick [nonfriend]	Stubbornness	Morality	Mood	Friendship
ask Food Taken	0	-0.5	-0.5	0
ask HP	0	-0.5	-0.5	0
ask Intended Food Intake	0	0	0	0
Request Take Food	1	-0.5	-1	-1
Request Leave Food	1	-0.5	-1	0

HP inc last tick [nonfriend]	Stubbornness	Morality	Mood	Friendship
ask Food Taken	0	0	0	-1
ask HP	-1	1	1.5	1
ask Intended Food Intake	0	0	0	0
Request Take Food	0	0	0	0
Request Leave Food	-1	1	1.5	1

Sum total	Stubbornness	Morality	Mood	Friendship
Request Leave Food	0	0	0	0

This zero-sum agent is how our agent is currently set, it represents the neutral agent. One who possesses no inherent emotional bias. This was done by balancing the positive and negative reactions to a friend, and those to a non-friend separately. We noticed the need for this when our agent would initially get stuck at the +ve and -ve limits of certain emotions. As future work, it would be exciting to see the emergent strategies which we hope to observe as we introduce biases. For instance, an agent whose net sum of all interactions leaves them with a stubbornness increase, coupled with an inherently moral agent, may create strategies yet thought of.

6.2.5 Proposing a Treaty

When the agent HP is critical, they will ask for help by either asking for the person upstairs to leave at least the amount of food necessary to survive or, if it doesn't have any active treaties, a treaty that states: if HP is higher than 20 leave 95% of the food in the platform for 3 days. At this time the agent's self-preservation instinct takes precedence and their objective becomes to eat enough food to get out of the critical state before the agent dies. When the agent is not in critical condition, it will be able to propose other treaties depending on their morality. The first treaty is detrimental for the tower and the other two are 'fair' treaties for the tower.

1. If the agent is very immoral, it will send a treaty that is detrimental for anyone above them. It is expected for this treaty to be rejected by any sensible agent's and is done as a way of mimicking frustration or the desire to hurt others when morality is low.
2. If HP is below 60 leave at least 60% of the food on the platform for 5 days. This treaty tries to encourage immoral agents in higher positions to share food.

3. If HP is below 10 leave at least 95% of the food on the platform for 10 days. This treaty is aimed at moral agents that would be willing to start eating less until they are close to death.

Both of the last treaties are proposed randomly by the agent, without a specific proclivity for any of the two. The randomness of the treaty selection is done to increase the volitional behaviour of the agent. Characteristics that the agent inherently lacks if its personality is solely based on the TRA.

6.2.6 Handling a Proposed Treaty

Following our strategy, agents that are in a bad mood will not accept treaties that they need to honour when they are in very low health, since they have low confidence in the capabilities of other agents. However, when in a good mood agents will accept treaties that must be honoured even at low health. Morality is an indicator of the agent’s desire to perform actions that will benefit others, as opposed to just benefiting the agent itself. Therefore, very moral agents will accept treaties that involve them consuming just enough food to survive, but not too little food that they die or too much that they are too greedy. Similarly, very immoral agents will only accept treaties where they consume more than enough food since they disregard the food needs of the other agents. The table below displays the mood and morality values that agents must satisfy to accept treaties with any given condition depending on the agent’s health, and the food the agent would have to leave according to the request of the treaty.

mood (m)/morality (M)		0 < morality < 40	20 < morality < 60	40 < morality < 80	60 < morality < 100	morality > 80	morality > 100
	AgentPosition \ FoodTaken	VeryLarge (1)	Large (2)	Moderate (3)	Little (4)	SurvivalAmount (5)	TooLittle (6)
mood > 0	Strong (1)	m>0, 0<M<40	m>0, 20<M<60	m>0, 40<M<80	m>0, 60<M<100	m>0, 80<M	N/A since 100<M
mood > 20	Healthy (2)	m>20, 0<M<40	m>20, 20<M<60	m>20, 40<M<80	m>20, 60<M<100	m>20, 80<M	N/A since 100<M
mood > 40	Average (3)	m>40, 0<M<40	m>40, 20<M<60	m>40, 40<M<80	m>40, 60<M<100	m>40, 80<M	N/A since 100<M
mood > 60	Weak (4)	m>60, 0<M<40	m>60, 20<M<60	m>60, 40<M<80	m>60, 60<M<100	m>60, 80<M	N/A since 100<M
mood > 80	SurvivalLevel (5)	m>80, 0<M<40	m>80, 20<M<60	m>80, 40<M<80	m>80, 60<M<100	m>80, 80<M	N/A since 100<M

Treaties also have a specified duration and signature count, which are visible to any agent that wishes to sign the treaty. These are likely to affect an agent’s decision to accept a treaty and have been taken into account. If a treaty lasts longer than twice the reshuffle period, an agent deems it too far into the future to be able to predict what will happen and plans to reject the treaty. However, if a treaty has more than five signatures, we deem it to be a popular treaty. If the agent plans to reject the treaty as a result of any previous logic, the influence of peer pressure will cause the agent to reconsider accepting the treaty. Peer pressure is based on the concept of crowd suggestibility coined by Park, R. E. in 1904 in his “Masse und Publikum”, Park [1972]. The stubbornness of an agent is used as the likelihood to give in to peer pressure, so more stubborn agents are less likely to reconsider their response, and more likely to stick with their original decision.

6.3 Food Intake

For an agent to choose what to eat on a given day, they first take into account treaties to which they have agreed. These treaties take the highest priority when deciding how much food to take from the platform. Therefore, when the agent in question has accepted a treaty the choice becomes dependent on the treaty conditions.

For each treaty that has been accepted, the agent follows the following logic:

1. They start by checking whether the condition stated within the specific treaty is satisfied, in which case the agreed-upon amount of food to leave on the platform is calculated.
2. In the case where the treaty request is not to leave food on the platform but to inform the sender of the treaty of our agent's HP, a message containing our current HP is sent to the sender agent.

After all the treaty conditions have been dealt with the agent calculates how much food to eat itself. The calculation depends on the morality of the agent, how much food it has decided to leave on the platform and how much food it has decided to eat. It must be noted that during other processes external to the final decision of how much to eat the agent might have decided how much it would like to eat, a non-binding decision.

Depending on the agents' decision regarding how much food to eat and leave it will arrive at 4 possible scenarios:

1. If it has decided how much food to eat and leave and both conditions can be met the agent will eat. If meeting both conditions is impossible, there is not enough food on the platform, a decision that fulfils the value of how much food to leave will be taken, with the value of how much the agent eats ranging from all that it can to just what it needs to survive.
2. If the agent knows how much it wants to eat but not how much to leave it will attempt to eat the amount it has decided it wants.
3. If the agent has not decided how much to eat but knows how much to leave it will eat depending on its health condition. If its health is critical or about to die it will eat whatever it needs to survive. If its health is not of concern it will eat as much as it can while leaving the amount it has decided to leave. Morality is used to scale how much food to eat.
4. If it is unaware of how much to leave or eat its decision will depend on its health level. If its health is critical or about to die it will eat whatever it needs to survive. If its health is not of concern, eat an amount of food scaled by its morality.

Chapter 7

Team 4 Agent Design

Chapter 8

Team 5 Agent Design

Chapter 9

Team 6 Agent Design

Chapter 10

Team 7 Agent Design

Chapter 11

Experiments

Experimental design - What are we looking to find – Conditions that lead to stability/instability
- Independent and dependent variables – Experimental parameters - Measurement methods

Chapter 12

Results

Chapter 13

Discussion

Chapter 14

Conclusion

Conclusion section.

Chapter 15

Future Work

Futur Work section.

Appendix A

Appendix

Appendix Section. Org chart could go here. Potentially could include implementation details.

Bibliography

- V. Ocaya, "Causal explanation in relation to human behaviour and moral responsibility," *Unknown*, 1985. [Online]. Available: <https://heinonline.org/HOL/LandingPage?handle=hein.journals/zambia17&div=5&id=&page=>
- D. M. I. S. Norazlan Hasbullah, Prof. Dr. Abdul Jumaat Mahajar, "A conceptual framework of extending the theory of planned behavior: The role of service quality and trust in the consumer cooperatives," *International Journal of Business and Social Science*, 2014.
- M. Fishbein, *Reasoned Action, Theory of*. John Wiley and Sons, Ltd, 2008. [Online]. Available: <https://onlinelibrary.wiley.com/doi/abs/10.1002/9781405186407.wbiecr017>
- C. D. Batson, J. G. Batson, R. M. Todd, B. H. Brummett, L. L. Shaw, and C. M. R. Aldeguer, "Empathy and the collective good: Caring for one of the others in a social dilemma." *Journal of Personality and Social Psychology*, vol. 68, no. 4, p. 619–631, 1995. [Online]. Available: <https://psycnet.apa.org/record/1995-25033-001>
- R. Shao, K. Aquino, and D. Freeman, "Beyond moral reasoning: A review of moral identity research and its implications for business ethics," *Business Ethics Quarterly*, vol. 18, no. 4, p. 513–540, 2008.
- A. Blasi, "Moral cognition and moral action: A theoretical perspective," *Developmental Review*, vol. 3, no. 2, p. 178–210, 1983.
- F. Wijermans, "Understanding crowd behaviour: simulating situated individuals," Ph.D. dissertation, 2011, relation: <http://www.rug.nl/> Rights: University of Groningen.
- R. Sokolowski, "Phenomenology of friendship," *The Review of Metaphysics*, vol. 55, no. 3, pp. 451–470, 2002. [Online]. Available: <http://www.jstor.org/stable/20131748>
- M. E. McCullough, M. B. Kimeldorf, and A. D. Cohen, "An adaptation for altruism: The social causes, social effects, and social evolution of gratitude," *Current Directions in Psychological Science*, vol. 17, no. 4, pp. 281–285, 2008. [Online]. Available: <https://doi.org/10.1111/j.1467-8721.2008.00590.x>
- P. Ekman, "An argument for basic emotions," *Cognition and Emotion*, vol. 6, no. 3-4, p. 169–200, 1992.
- D. Kahneman and A. Tversky, *Choices, Values, and Frames*, p. 1–16, 2000.
- K. Frankish, "Dual-process and dual-system theories of reasoning," *Philosophy Compass*, vol. 5, no. 10, p. 914–926, 2010.
- P. Kamble, "What is subconscious mind? how does it impact our behaviour?" *SSRN Electronic Journal*, 2021.

.-. Park, Robert Ezra, *The crowd and the public, and other essays.*, ser. Heritage of sociology. Chicago,: University of Chicago Press, 1972, "The crowd and the public [Masse und Publikum] translated by Charlotte Elsner. Note on The crowd and the public by Donald N. Levine."