Distributed Social Multi-Agent Negotiation Framework For Incomplete Information Games

Joseph Walton-Rivers[*], Edward Longford[†], Daniel Gomme[‡], Richard Bartle[§], Michael Gardner[¶]

School of Computer Science and Electronic Engineering
University of Essex, Colchester
Colchester, UK
[*]jwalto, [†]elongf, [‡]dsdgom, [§]rabartle, [¶]mgardener@essex.ac.uk

Abstract—In this paper, we propose a social negotiation system in which agents can communicate and interact with each other socially throughout a *Sheriff of Nottingham* game. We address issues with the number of options available while negotiating, particularly when bluffing is involved. Experiments are proposed that would allow us to validate how closely this framework mirrors real social interaction in the game, and the possibility of generalising multi-agent negotiation beyond this framework is raised.

Index Terms—social, agents, framework, communication, negotiation, game, interaction

I. INTRODUCTION

Interest has been growing in enhancing cooperation strategies for distributed AI agents in dynamic (non-fixed) teams with either unstructured or minimally structured communication. This paper sets out plans for a communication framework in a modified version of *Sheriff of Nottingham* [1], allowing agents to participate in co-operation and adversity with other players via negotiation. We combine this with the Mod-SYMLOG social model of behaviour [2] to take into account perceptions of other players, with the aim of creating AI agents that interact in a more socially believable manner. Such agents may significantly improve player experience, as well as open up opportunities for other games to adopt similar techniques.

Sheriff of Nottingham is a suitable first testing ground. Its negotiation involves bluffing and rewards, providing an opportunity for players to participate in complex co-operation and deception, and to interact with each other socially through these.

II. BACKGROUND

A. Communication

Communication can be defined as either closed query/return or open query/return [3], which this paper shall refer to as Structured and Unstructured Communication respectively.

Structured communication is a series of predefined queries with a set of predefined returns as answers. One example of structured communication within games is that of the card game *Bridge*. During the "bidding" phase of gameplay, each player communicates to their "partner" information about the cards they hold. These signals sent to partners are predefined within the game rules and supply a range of information depending on the query and return between partners. I.E.

players are limited to a range of predefined queries and/or statements they can make about their cards [4].

Unstructured communication has no defined query/return. This is often how humans interact, through a free flow of information rather than following a strict structure. In games this can take the form of negotiation between 2 or more players, where 1 or more players try to persuade another player(s) to cooperate to achieve a certain goal. Games such as *Settlers of Catan* have negotiation phases where players will partake in unstructured communication in order to exchange in-game resources [5].

An additional factor in communication is the level of information that an AI agent is exposed to when it makes a decision. For example, players do not know the value of cards in other players' hands in *Poker*, or *Hearthstone*. That information is hidden from them, and therefore the agents are working with incomplete information.

The uncertainty introduced causes issues for agents that look into the future, such as Monte Carlo Tree Search [6] and minimax [7]. There are now far more possibilities, many of which are untrue, so finding 'good' results takes much longer.

B. Social Multi-agent

Multi-agent Domains involve 2 or more agents interacting. These agents can either be centralised, where all agents are controlled by a central controller, or distributed, where each agent has its own controller passing instructions. Focusing on Distributed agents for negotiation allows for independent controllers to either cooperate or compete based on Incomplete Information Sharing.

To enhance the AI agents decision-making skills, a social model of other players can be constructed. This can be done as a function of trustworthiness of an another agent or assigning personality types. For this Framework, the agents will be utilising a Modified version of the SYMLOG Social Model. This Modification rates each agent as a value on 3 axes. 1) Whether the other agent is being Dominate or Submissive (UD) towards the Rating Player. 2) How Positive or Negative (PN) the other agent is behaving towards the rating agent. 3) If the other agent is working towards or against (FB) the goals of the rating agent [2].

C. Cooperative Multi-agent Games with Incomplete information currently under research

1) Diplomacy: Perhaps one of the most complicated negotiation games to model is that of *Diplomacy*. *Diplomacy* is an imperfect information board game of up to 7 players. Moves in the game are made simultaneously and each player can make multiple public and private nonbinding and contradictory agreements through unstructured communication between any number of the other players, with dynamic factions of unequal sizes, and allowing only 1 winner in the standard rule set [8].

While computer-based versions of *Diplomacy* have been available since 1984 [9], the AI has often lacked the negotiation skills required to play against human players and often limited the types of agreements to a more structured communication approach. Efforts to create an AI capable of complex unstructured negotiation are ongoing [10], including the development of a communication language to allow players to perform complex negotiations [11]. AI algorithms used so far have included: de Jonge's Branch & Bound search and Ferreira's DipBlue trust based modelling of other players [10]

2) Hanabi: *Hanabi* is a game for 2 to 5 players, all of whom are acting as one team and it is the team that wins, as opposed to *Diplomacy* where only one player can win. Each player is dealt 5 cards. *Hanabi* varies from other negotiation card games, as the player is the only person who is unaware of their cards. The group is assigned a limited number of tokens, which are used to provide information to other players. The objective of the game is to complete the 5 "suits", which have assigned number values of between 1 and 5. The game is lost if the players run out of life tokens or cards [12].

III. PROPOSED EXPERIMENTAL SETUP

We implement *Sheriff of Nottingham*[1], and a communication framework agents can use to play and thereby interact socially.

A. Game Description

The game consists of:
- 216 Cards - 114 legal Goods, 60 illegal Goods, 12 Royal Goods
- 110 Gold Coins - 39 1-gold coins, 42 5-gold coins, 17 20-gold coins, 12 50-gold coins
- 5 Merchant bags
- 5 Merchant Stands
- 1 Sheriff Marker

1) Initial Set up: Each player is allocated 50 gold worth of coins (1 x 20-gold, 5 x 5-gold, 5 x 1-gold), allocated a Merchant Stand, a Merchant bag, and dealt 6 cards from the deck. One Player is then selected to play the role of Sheriff for the first round of gameplay. Two "discard piles" are then created next to the deck of cards by removing the top 10 cards from the deck and placing half on either side, face up, allowing all players to observe the discarded cards.

2) Game Play: The objective of *Sheriff of Nottingham* is to move the highest value of goods past the Sheriff and onto your Merchant Stand. Cards are moved, via each Merchant's bag (thus no other player having information over which cards are being moved) from a player's hand to their Stand. Merchant Players are required to negotiate with the Sheriff for their bag not to be inspected during this process. Each Game turn is described in more detail below.

- Phase 1: Market

All Merchant Players then are given the opportunity to alter their cards. Starting with the Merchant Player to the sheriff's left and continuing in a clockwise direction, each Merchant Player can choose to discard up to 5 cards and draw the corresponding number from the deck. These cards can either be taken publicly from 1 of the discard piles, secretly from the deck, or a combination of the two.

- Phase 2: Load Merchant Bag

Once all Merchant Players have had the opportunity to exchange cards, each Merchant Player secretly chooses between 1 and 5 cards and places them within their bag. Once the cards have been placed into the bag the Merchant Player cannot change their minds. At this point, the only information any player has is the cards in their own hand, the cards they have placed in the bag (if they are a Merchant), and the visible cards on the discard pile.

- Phase 3: Declaration

Each Merchant Player then declares to the Sheriff the number of cards they have in their bag (this must be honest) and then what type of card they are. The player can only declare one type of card. For example a player can place into their bag 3 cards consisting of 2 apples and 1 bread but they can only declare one of these two goods, so they may say "this bag contains 3 apples". Players can only declare legal goods.

- Phase 4: Inspection (and Negotiation)

When all Merchant Players have declared what the contents of their bag is, the Sheriff Player is given the opportunity to inspect these bags. Before the bag is opened, each Merchant Player has the opportunity to negotiate with the Sheriff not to open their bag. Merchant Players can attempt to bribe the Sheriff with gold, cards within their bag, cards on their stand, future favours, or a combination of those options for the Sheriff not to open their bag. Once the initial offer has been placed the Sheriff can choose to accept the offer, reject it or negotiate an amendment. Other Merchant Players can also join the negotiation to persuade the Sheriff Player to open the opponent's bag with their own bribes.

This negotiation strategy is the main focus of our paper. We shall attempt to provide a mechanism for the AI agents to create offers and counteroffers based on known cards played and Social Models of the other AI agents

If the Sheriff Player allows the Merchant Players bag to pass inspection, the Merchant Player then opens their bag and displays the legal goods on their Merchant Stand. Illegal goods are kept face down, preventing other players from knowing what illegal goods have been passed. If the Sheriff Player

Goods	Value	King bonus	Queen bonus
Legal Goods			
Apples	2	20	10
Cheese	3	15	10
Bread	3	10	5
Chickens	4	10	5
Illegal Goods			
Pepper	6	-	-
Mead	7	-	-
Silk	8	-	-
Crossbow	9	-	-

TABLE I
TYPES OF GOODS AND THEIR VALUES

chooses to inspect the bag, it is opened publicly and displayed to all other players. If the Merchant Player is honest, the Sheriff Player pays a penalty to the Merchant Player whose bag has been opened. If the Player is found to be lying, they need to pay the Sheriff Player a fine for all undeclared goods, 2-gold for undeclared legal goods, 4-gold for each undeclared Illegal good, and between 3 and 5-gold for Royal goods. The goods which were not declared are discarded into the discard pile. Any goods that were declared will move to the Merchant Players stall regardless.

- Next Turn

The role of Sheriff is then moved to the next player (in a clockwise motion) and the above phases are repeated. The game continues until all Players have been Sheriff twice. Once the game ends, all cards that are not on each of the Player's Merchant Stall are discarded.

3) Scoring: Each of the 216 cards are provided with monetary value, based on the type of goods each card represents (see Table I), a penalty value if the card is found on inspection and bonus points for a player having either the most of that type of cards at the end of the game (King Bonus) or Second (Queen Bonus). Royal Goods add an additional number of a good type to a Players Merchant Stand before calculating which player is King or Queen of that particular good type.

The players score is the sum of value of their gold, the face value of goods cards in their Merchant Stand, king bonus(es) if any, and queen bonus(es) if any.

IV. COMMUNICATION

A. AI Obstacles

The idea of offering the sheriff a bribe poses a significant difficulty to any form of AI affected by how large an action space is. There are an immense number of potential bribes that may be offered, as the player may offer goods from their bag that the bag does not actually contain. They may also offer any combination of goods from their stall, as well as any amount of gold under the total gold that they own.

For example, take a player with a bag containing 5 goods, currently in negotiation with the sheriff. They may offer up to 5 goods to the sheriff from their bag, none of which have to be present in the bag. Assuming 4 players playing, with no Royal Goods, there are a total of 8 possible goods. The

player may also offer 'contraband', in the more general sense. So, choosing 5 (k) times from 9 (n) possibilities, with order being unimportant and repetition allowed, we end up with a total of $1,287$ possibilities, for if the player offers 5 goods. However, players may offer between 5 goods and 1 good. This results in $2,001$ possible offers for a single bag.

$$\sum_{k=1}^{5} \binom{n+k-1}{k} = 2001, \text{ where } n = 9$$

Such a large branching factor makes it incredibly difficult to search the state space fully - for example, MCTS can be hobbled by high branching factors[13][14]. Therefore, we need to refine bribe options to make it much easier for agents to navigate, but in a way that does not lose core functionality of the negotiation.

However, before we restructure the bribery offers, we must lay out how negotiations occur.

B. Negotiation Strategy

The negotiation is structured to allow each agent a chance to participate in the negotiations. Firstly, the sheriff selects one of the players whose bag has not been opened and prompts them for an offer. The owner of the bag may then present an offer to the sheriff. All other players then get the option of proposing a counteroffer for the sheriff to open the bag. The sheriff is then given the option of either accepting any of the offers presented to them, or returning to the owner of the bag with a counter offer. The negotiation proceeds as such until either the sheriff accepts an offer from one of the players, or chooses to accept no offer and either allow the owner of the bag to pass through anyway, or open the bag. Counteroffers may build upon previous offers by adding another offer to a previous one.

C. Offer Options

The possible bribery actions can be split into two distinct roles, firstly, the owner of the bag negotiating deals for allowing the player through without the bag being opened. In contrast to this, other players are making deals to entice the sheriff to open the bag. Deals that are made by other players cannot include offers of the owner's bag contents, as they are not the bag owner. This includes any additions a Sheriff makes to an offer when countering.

We must also remove the ability to use arbitrary pretences of bag contents for negotiation, as this increases the number of possible offers, and therefore the action space, by a great deal. Similarly, offers exchanging varying numbers of an item (1/2/3.. of X) also increase the number of possibilities.

For arbitrary bag content pretences, we use random generation that stays fixed through the rest of this bag's offer process. For numbers, we instead use fixed ratios of other information available, including from generated false information. We take care to provide, for each bribe depending on hidden information, a version depending on true information, and a version depending on false information (a lie). We also make

sure that these versions are indistinguishable to agents that do not know the contents of the bag.

The offers provided also cover every possible kind of token the merchant has license to offer - stall goods, bag goods, gold, and gold dependent on bag goods or declared goods.

The bribes that an agent Merchant may offer the Sheriff, in order to entice the sheriff to either inspect the bag or allow it to pass, are as follows:

- Dependent on declared goods
  - Half declared value of goods (in coins)
  - Half declared goods in bag (rounded down)
  - Full penalty value of declared goods
- Bag items
  - Truth - Half contraband in bag (rounded down) (minumum 1)
  - Truth - Half undeclared legal goods in bag (rounded down) (minimum 1)
  - Lie - 2 randomly generated contraband items, 1 if the bag size is 3 or less
  - Lie - 2 randomly generated legal goods (rounded down) (minimum 1)
- Independent token offers
  - 2 Gold (based on observed play as an offer)
  - 1 good from stall (cheese, chicken, bread, apples)

These actions combined allow for putting together offers of varying value and level of truth, allowing for rich interaction with the social model.

D. Mod-SYMLOG interaction and application

The above communication framework means the game's communication is structured. This allows mapping human and agent actions, including their communications, into the Mod-SYMLOG social model [2]. Agents can evaluate where players lie within the Mod-SYMLOG space, using metrics such as how often unenforced deals were actually followed through with, the tendency to work towards common goals, how much another player's lies caused loss, and so forth.

Once a player has been placed into the model, that information can then be used to influence future decisions, and to create agents that elicit certain perceptions from other players, possibly using them to their advantage and allowing for more believable play.

V. PROPOSED EXPERIMENTS

It would be useful to verify that the agents' reduced offer space adequately interacts with the social model. Therefore, humans should play several games under the agents' offer restrictions, recording their perceptions of other players using the Mod-SYMLOG social model [2] at the end of each round. They should then do the same while playing the physical version of the board game, unrestricted. If there are differences, we may analyse them to find weaknesses in our agents' offer space.

Once interaction with the social model is verified, we can create agents to play alongside humans in this framework,

some aware of the social model, and some not. We can then see if humans prefer playing against the socially-aware agents, and whether they find them more 'believable'.

VI. CONCLUSIONS AND FUTURE WORK

This paper describes one of the main technical limitations for this kind of game - the need for a structured communication channel for the agents. It proposes a communication system that addresses this, both providing a structure and alleviating the issues with the otherwise very large number of possibilities while communicating. It lays out the groundwork for future experiments using the communication system described in the paper.

There is room for improvement with the process of negotiation. It may be possible to expand it to a form of general multi-party negotiation with flexible rules, and previous negotiations themselves being considered in the process.

REFERENCES

[1] S. Halaban and A. Zatz, "Sheriff of Nottingham," 2014.
[2] E. Longford, M. Gardner, and V. Callaghan, "Social Organisation and Cooperative Learning: Identification and Categorisation of Groups and Sub-Groups in Non-Cooperative Games," in *Immersive Learning Research Network* (D. Beck, A. Peña-Rios, T. Ogle, D. Economou, M. Mentzelopoulos, L. Morgado, C. Eckhardt, J. Pirker, R. Koitz-Hristov, J. Richter, C. Gütl, and M. Gardner, eds.), (Cham), pp. 131–143, Springer International Publishing, 2019.
[3] S. Dolnicar, "Asking good survey questions," *Journal of Travel Research*, vol. 52, no. 5, pp. 551–574, 2013.
[4] "Standard English Modern Acol System File," 2014.
[5] K. Teuber and G. R, "Catan 5Th Ed Rules English," 2015.
[6] G. Chaslot, S. Bakkes, I. Szita, and P. Spronck, "Monte-carlo tree search: A new framework for game ai.," in *AIIDE*, 2008.
[7] B. W. Ballard, "The*-minimax search procedure for trees containing chance nodes," *Artificial Intelligence*, vol. 21, no. 3, pp. 327–350, 1983.
[8] A. B. Calhamer, "The Rules of Diplomacy," 2000.
[9] A. Hill, "Computer Diplomacy," 1984.
[10] J. Marinheiro and H. Lopes Cardoso, "Towards General Cooperative Game Playing," in *Transactions on Computational Collective Intelligence XXVIII* (N. T. Nguyen, R. Kowalczyk, J. van den Herik, A. P. Rocha, and J. Filipe, eds.), (Cham), pp. 164–192, Springer International Publishing, 2018.
[11] A. Fabregues and C. Sierra, "DipGame: A challenging negotiation testbed," *Engineering Applications of Artificial Intelligence*, vol. 24, no. 7, pp. 1137–1146, 2011.
[12] J. Walton-Rivers, P. R. Williams, R. Bartle, D. Perez-Liebana, and S. M. Lucas, "Evaluating and modelling hanabi-playing agents," in *2017 IEEE Congress on Evolutionary Computation (CEC)*, pp. 1382–1389, IEEE, 2017.
[13] C. B. Browne, E. Powley, D. Whitehouse, S. M. Lucas, P. I. Cowling, P. Rohlfshagen, S. Tavener, D. Perez, S. Samothrakis, and S. Colton, "A survey of monte carlo tree search methods," *IEEE Transactions on Computational Intelligence and AI in games*, vol. 4, no. 1, pp. 1–43, 2012.
[14] H. Finnsson and Y. Bjornsson, "Game-Tree Properties and MCTS Performance," in *International Joint Conference on Artificial Intelligence*, p. 8, 2011.