1

Introduction

Ants and bees can also work together in huge numbers, but they do so in a very rigid manner and only with close relatives. Wolves and chimpanzees cooperate far more flexibly than ants, but they can do so only with small numbers of other individuals that they know intimately. Sapiens can cooperate in extremely flexible ways with countless numbers of strangers. That's why Sapiens rule the world, whereas ants eat our leftovers and chimps are locked up in zoos and research laboratories.

Yuval Noah Harari, *Sapiens* (2015)

1

We rule the world. In other words, the human species is the most dominant species on earth (according to us humans). Yuval Harari [65] stated that this could be attributed to our unique ability to cooperate both flexibly and with very large numbers. Crucial to our cooperative behaviour is our ability to communicate and, as a special case of communication, negotiate [36, 105]. Negotiation enables humans to resolve conflicts over resources, improving collective productivity through task specialisation. For instance, a hunter can exchange surplus food for a blacksmith's traps, allowing each to focus on their expertise. Now, the hunter can keep hunting, and the smith can keep smithing instead of both having to do both. This mutually beneficial agreement, facilitated by negotiation, improves the overall efficiency and productivity of a community. Humans are not the only species that negotiate, as studies have shown that, e.g., chimpanzees also possess this ability [105]. However, human communication and negotiation are far more complex and likely play a crucial role in our dominant position.

Negotiation is an important part of human interaction in modern society, from interpersonal relationships to international diplomacy. It serves as a critical mechanism for resolving conflicts, allocating resources, and enabling cooperation between parties with divergent interests. Negotiations are so embedded into our daily lives that we do not always notice them [117]. Discussing the details of a contract is clearly a negotiation, but other negotiations are not that explicit. For example, finding a meeting slot with colleagues, deciding where to go on a city trip with a friend, or navigating busy bike lanes with non-verbal signals can all be considered forms of negotiation.

Negotiation has seen interest from the scientific community, for example, by social scientists [131, 149], economists [118, 115] and by mathmaticians [111, 132]. With the advent of artificial intelligence, it is envisioned that artificial intelligent agents also have conflicts of interest, which should be studied. This dissertation, therefore, studies negotiation in computer science, more specifically, in multi-agent systems.

1.1 NEGOTIATION IN MULTI-AGENT SYSTEMS

As artificial intelligence becomes more embedded into society and daily life, the question arises of how autonomous or semi-autonomous agents interact. This is a central focus of Multi-Agent Systems (MAS), a broad research area within AI concerned with "systems that include multiple autonomous entities with either diverging information or diverging interests, or both" [140]. MAS research encompasses a range of research areas, such as multi-agent learning [151], communication and consensus [121], organisational structures [72], and distributed constraint satisfaction [164].

In multi-agent interactions, situations arise where agents, whether AI-AI or AI-human, must cooperate despite (partially) conflicting interests. Such conflicts of interest should then be resolved to improve payoff or even obtain payoff in the first place (e.g., surveying drone swarms or transporting goods using multiple robots). Resolving the conflicting interests can be done in several ways. In some cases, conflicts are handled through adherence to pre-defined institutions or social norms (e.g., traffic rules governing right-of-way). In others, predefined coordination mechanisms or protocols suffice for enabling coordination.

If there is no predefined method of handling conflicts of interest, then negotiation or bargaining can be used to enable cooperation. As it was important for humans to develop negotiation skills, AI researchers envisioned that this is also an important skill for intelligent agents to resolve conflicting interests. This led to the research field of automated negotiation [79]. Automated negotiation plays a role in real-world applications, such as traffic light coordination [62], calendar scheduling [125], or balancing energy demand and production in local power grids, and also in games, such as Diplomacy [107] and Werewolves.

Automated negotiation is a multidisciplinary area at the intersection of artificial intelligence, game theory, and decision science. Early work in this field, pioneered by scholars such as Smith [143], Rosenschein [130], Klein and Lu [88], and Sycara [147], laid the groundwork for computational approaches to negotiation. The field saw significant advancements with the development of more sophisticated negotiation strategies and frameworks. Notably, the work of Faratin et al. [48] introduced time-dependent and behaviour-dependent tactics, which have become fundamental components of many negotiation strategies.

Over the years, the research community has developed a wide array of negotiation strategies, protocols, and evaluation frameworks, leading up to initiatives like the Automated Negotiating Agents Competition (ANAC) [16] and the General Environment for Negotiation with Intelligent multi-purpose Usage Simulation (GENIUS) negotiation platform [99]. The combined effort of GENIUS and ANAC provided a standardised test bed with more than 100 negotiating agents and negotiation scenarios that are readily accessible for research on automated negotiation [9]. Both the development of new agents and structured evaluation of these agents are important, which we will discuss in Sections 1.1.1 and 1.1.2, respectively.

1.1.1 Developing negotiation agents

The goal of automated negotiation research is the development of agents that are capable of negotiating quickly and effectively. The negotiating agents are generally hard-coded strategy algorithms with parameters that are tuned during design to optimise their performance. Such agents must navigate large outcome spaces, deal with incomplete information, and engage in strategic interactions with other parties, all while attempting to achieve optimal results for themselves or the humans or organisations they represent. Traditionally, developing negotiation agents relied on manually designed heuristics and strategies, optimised and tested on constrained sets of negotiation settings, to make it manageable. This is still a commonly seen approach in recent editions of ANAC [5]. While such methods have proven effective in specific scenarios, they struggle to generalise across diverse negotiation scenarios and opponent types. This limitation becomes problematic as we envision deploying automated negotiation agents in real-world applications, where they may encounter a wide variety of negotiation problems and counterparts.

To alleviate the difficulty in designing negotiation agents manually and to obtain better performance, optimisation methods were later used to (partially) automate the design. This allows agents to be reconfigured in various negotiation settings, improving generalisability. Genetic algorithms have been used, but optimising negotiation agents is computationally expensive, due to, e.g., searching for potential agreements in outcome spaces exponential in the objectives for which a decision must be made. To circumvent the computational complexity issue, agents were configured on small sets of scenarios and opponent types [104]. For instance, agents were only tested in one or two scenarios [71] or merely optimised against themselves [46, 43]. The resulting agents are highly specialised and have unpredictable performance when negotiating "out-of-distribution".

The problem we address is the development of agents capable of learning to negotiate effectively across large and diverse sets of negotiation settings. Learning to negotiate involves agents optimising policies for actions, such as making proposals, accepting offers, or walking away, to maximise their performance. We adopt a learning-based approach because hand-designed negotiation agents impose human-induced bias, leading to suboptimal performance, limitations we want to overcome. Furthermore, as traditional hand-crafted negotiation state representations used for agent input can discard informational value, we also aim to learn representations directly from (near-) raw observational data.

Part one of this dissertation aims to address the problem of learning autonomous agents to negotiate on large and diverse sets of opponent types and negotiation scenarios. We explore approaches to learn negotiation strategies more efficiently (Chapter 3). We also explore strategy-switching mechanisms that are learned based on characteristics of a negotiation setting in an effort to further improve the adaptiveness of negotiation strategies (Chapter 4). Finally, we leverage machine learning, more specifically reinforcement learning, to push the boundaries of what is possible in automated negotiation (Chapter 5). We discuss all three approaches briefly below.

ALGORITHM CONFIGURATION

Chapter 3 focuses on enhancing existing negotiation strategies through automated algorithm configuration. Many negotiation agents in the literature use strategies with fixed parameters or ones manually tuned by their designers. This approach, while straightforward, often leads to suboptimal performance and poor generalisation across different negotiation settings.

We propose to address this limitation by applying general-purpose automated algorithm configuration techniques, specifically Sequential Model-based Algorithm Configuration (SMAC) [75], to the task of optimising negotiation strategies. SMAC is a generally well-performing algorithm that is often used for such tasks. Our approach automatically configures the parameters of commonly used negotiation strategy components, including those related to bidding strategies, accepting strategies, and opponent modelling. Our approach employs a set of features that characterise negotiation scenarios and opponent behaviours, enabling more efficient configuration across diverse negotiation settings. This enhanced efficiency enables us to optimise performance on larger and more diverse sets of negotiation simulations than previously achievable in automated negotiation systems. We have been the first to pursue this approach, which goes beyond simple application of algorithm configuration.

ALGORITHM SELECTION

Building on the insights gained from automated configuration, we next explore the potential of portfolio-based methods in automated negotiation in Chapter 4. This direction is motivated by the observation that there is often no single best strategy for all negotiation settings — different strategies may be more or less effective, depending on the specific negotiation scenario and opponent type[99].

Our portfolio construction method, based on Hydra [162], iteratively generates complementary strategies that specialise on sub-spaces of the negotiation game distribution. We build on top of the previous work where we configure a single-best strategy and, instead, now configure a portfolio of complementary strategies. We then employ a learned strategy selection model, using AutoFolio [100], that selects a strategy from the portfolio for a given negotiation game. This portfolio-based approach offers the advantage of better adaptation to different opponents and negotiation problems and outperforms the single-best strategy across diverse settings.

REINFORCEMENT LEARNING

In Chapter 5, we investigate the potential of end-to-end reinforcement learning approaches for automated negotiation. The previous algorithm configuration and selection approaches still rely on a high degree of manual design, as the parameterised strategies are a prerequisite for both approaches. This introduces bias, which can help in reducing the search space of negotiation strategies. However, it also potentially introduces suboptimality to the method. To further mitigate potential biases introduced by human design, we propose a conceptually straightforward reinforcement learning approach, based on Proximal Policy Optimisation (PPO) [135], in an effort to obtain better performance across a broader set of negotiation settings.

A key challenge in applying reinforcement learning to negotiation is handling the variability in problem structures. Negotiation outcome spaces are arbitrarily sized, leading to widely varying observation and action spaces. This variability poses difficulties for traditional neural network architectures used in reinforcement learning. To address this challenge, we developed a policy network using Graph Neural Networks (GNNs) [26] that operates on a custom graph-structured representation of negotiation game data and trained it through reinforcement learning. Our policy can process negotiation scenarios of varying sizes and structures, enabling generalisation across diverse scenarios. This approach opens up avenues for more sophisticated and adaptable negotiation agents in automated negotiation by leveraging reinforcement learning techniques. It significantly reduces human-induced biases in learned strategies while facilitating the integration of advanced learning capabilities into policy networks.

1.1.2 EVALUATING NEGOTIATION AGENTS

An essential part of developing negotiation agents is evaluating them empirically. Negotiation agents are tested by running computational experiments where agents compete against each other on sets of negotiation scenarios. Without a good performance measuring method, it is difficult to judge whether a designed strategy is effective or not. Over the years, multiple such evaluation methods were developed and tested on selected strategies [9]. Despite the efforts, there is still no clear answer to the question of what makes a negotiation agent good or not. To make matters more complex, we see an increase in the deployment of machine learning techniques in recently developed agents, which makes them change their behaviour over time. This inherent non-stationarity poses a challenge to empirical evaluation, as evaluation methods often require repeated trials under stable conditions. Evaluating learning agents via repetition yields non-stationary performance metrics, complicating such an evaluation. In part two of this dissertation, we study this in more depth.

EMPIRICAL ANALYSIS OF NEGOTIATION AGENTS

We organised ANAC 2021 and 2022, and set the challenge explicitly on the design of negotiation agents that learn and adapt their behaviour over time. The agents submitted to the 2022 competition are used for an extensive evaluation study of negotiation agents comparing various evaluation criteria in Chapter 6. We observe that multiple factors influence the performance of an agent. The evaluation criteria have an influence, but so does the composition of the group of opponents. We conclude that there is no single evaluation criterion to evaluate negotiation agents, and that outperforming a fixed group of opponents is no guarantee that an agent will perform well in general. However, a common practice in ANAC is to evaluate agents based on a single evaluation criterion in a fixed group of agents.

PROPOSED APPLICATION DOMAIN

We argue that the lack of a good performance measuring method goes hand in hand with the lack of a straightforward application domain. Different application domains likely require different performance criteria. Without a clear application domain, the community could be running in circles in pursuit of better strategies. To this extent, in Chapter 7, we propose an application domain for negotiation agents that we think is rich and useful. We propose the use of negotiation agents in calendar scheduling problems and map out the necessary steps to achieve this. We hope this sets an example and a direction for the automated negotiation community.

1.2 RESEARCH QUESTIONS

We formulate two main research questions and several sub-questions for this dissertation. The questions are in line with what was discussed in Section 1.1.1 and Section 1.1.2:

Q1 How do we design agents that can learn to negotiate?

SQ1.1 How can we reduce human-induced biases and conceptual complexity in learned negotiation strategies?

SQ1.2 Can we learn to generalise over diverse negotiation instances?

Q2 Is there a uniform way of evaluating negotiation agents?

SQ2.1 Is there a single-best metric?

SQ2.2 What is the value of the average utility metric for evaluating negotiation agents?

Investigating these questions will advance machine learning within automated negotiation, leading to more sophisticated negotiation agents while reducing design overhead. Furthermore, improved insight into evaluation metrics is essential for refining research goals and assessment methodologies. Progress in automated negotiation is important to stimulate the wider adoption of agent-based AI systems.

1.3 Dissertation structure

The remainder of this dissertation is structured as follows. Chapter 2 provides background information on commonly used concepts and topics covered in this dissertation. Chapter 3 presents our work on automated configuration of negotiation strategies. We detail our application of SMAC to strategy optimisation, which is not straightforward due to the combined nature of negotiation problems consisting of opponents and scenarios. We introduce our feature representation for negotiation problems and provide empirical results demonstrating the effectiveness of this approach. Chapter 4 describes our portfolio-based approach for strategy selection and adaptation. We explain our portfolio construction method, detail the AutoFolio-based strategy selection process, and present results from ANAC-like tournaments, showing the advantages of this approach. Chapter 5 introduces our end-to-end reinforcement learning method for general negotiation. We describe our graph-based policy architecture, explain the training process using PPO, and provide extensive evaluations against both baseline and state-of-the-art negotiation agents. Chapter 6 contains our extensive analysis of negotiation agents that were submitted to ANAC 2022. The 2022 edition focused specifically on learning negotiation agents. We compare various performance criteria and study the effect of non-stationary behaviour. We argue that evaluating negotiation agents is challenging and that there is no single best method to do so. Chapter 7 resurfaces a realistic real-world application domain for negotiation agents that has been studied in the past, namely, agent based calendar scheduling. We thoroughly dissect the complexity of the domain into manageable sub-problems to solve in the hope of setting the stage for future work in this application domain. Chapter 8 concludes the dissertation with a discussion of the broader implications of our work, its limitations, and promising directions for future research.