TECHNICAL UNIVERSITY OF CRETE

MASTER OF SCIENCE THESIS

Overlapping Coalition Formation under Uncertainty

Michail Mamakos

Supervisor:
Georgios Chalkiadakis

A thesis submitted in fulfillment of the requirements for the degree of Master of Science in

Electronic and Computer Engineering

awarded by the

School of Electrical and Computer Engineering

March 2017

Technical University of Crete

Abstract

School of Electrical and Computer Engineering

Master of Science

Overlapping Coalition Formation under Uncertainty

by Michail Mamakos

Research in cooperative games often assumes that agents have complete information regarding the coalitional values, and that they can belong to one coalition only. In this thesis, we remove these unrealistic restrictions, and study various aspects of uncertainty facing agents in coalition formation environments, while allowing them to belong to multiple coalitions simultaneously.

We begin by focusing on agent uncertainty regarding the resource contributions of potential partners. To tackle this, we provide three novel methods that obtain probability bounds for assessing the success of teams towards coalitional task completion. Our first method is based on an improvement of the Paley-Zygmund inequality, while the second and the third are devised based on the two-sided Chebyshev's inequality and the Hoeffding's inequality, respectively. Our methods allow agents to demand certain confidence levels regarding the resource contribution of coalitions; and agent beliefs are updated in a Bayesian manner, following formation decisions.

We then proceed to study situations where agent uncertainty is over the underlying collaboration structure, which determines the values of the (possibly overlapping) coalitions. In this context, we first propose a novel concise representation scheme, termed "Relational Rules", which extends the celebrated MC-nets representation to cooperative games with overlapping coalitions. We then present a novel decision-making method for decentralized overlapping coalition formation, which employs, for the first time in the coalition formation literature, "Probabilistic Topic Modeling" (a highly successful unsupervised learning approach). We demonstrate experimentally that by interpreting formed coalitions as documents, agents using our approach are able to effectively and efficiently learn profitable collaboration patterns (or "topics").

Abstract in Greek

Στα συνεργατικά παίγνια συχνά γίνεται η υπόθεση ότι οι πράκτορες έχουν πλήρη γνώση της χρησιμότητας που αποφέρει ο σχηματισμός των συνασπισμών, και ότι ο καθένας μπορεί να συμμετέχει μόνο σε ένα συνασπισμό. Στην παρούσα μεταπτυχιακή εργασία, αφαιρούμε αυτούς τους περιορισμούς, οι οποίοι συχνά δε συνάδουν με τα πραγματικά περιβάλλοντα. Ως εκ τούτου, μελετάμε διάφορες πηγές αβεβαιότητας με τις οποίες έρχονται αντιμέτωποι οι πράκτορες σε συνεργατικά περιβάλλοντα, ενώ τους επιτρέπουμε να συμμετέχουν σε πολλούς συνασπισμούς ταυτόχρονα.

Αρχικά εστιάζουμε στα προβλήματα που αντιμετωπίζουν οι πράκτορες σχετικά με την αβεβαιότητα που προέρχεται από τη συνεισφορά των εν δυνάμει συνεργατών τους. Για την αντιμετώπιση αυτού του προβλήματος αναπτύσσουμε τρεις μεθόδους οι οποίες βασίζονται στον υπολογισμό φραγμάτων πιθανοτήτων για την εκτίμηση της ικανότητας ομάδων πρακτόρων να πραγματοποιήσουν μία εργασία. Η πρώτη μέθοδος βασίζεται σε μία βελτίωση της ανισότητας Paley-Zygmund, ενώ η δεύτερη και η τρίτη βασίζονται στη διμερή ανισότητα του Chebyshev και την ανισότητα του Hoeffding, αντίστοιχα. Οι μέθοδοί μας επιτρέπουν στους πράκτορες να απαιτούν επίπεδα εμπιστοσύνης της επιλογής τους σχετικά με τη συνεισφορά πόρων των συνασπισμών. Οι πράκτορές μας διατηρούν Μπαεσιανές πεποιθήσεις, που ανανεώνονται μετά από κάθε σχηματισμό συνασπισμών.

Έπειτα, μελετάμε καταστάσεις στις οποίες οι πράκτορες έχουν αβεβαιότητα σχετικά με την υποκείμενη συνεργατική δομή, βάση της οποίας καθορίζονται τα κέρδη των (πιθανώς επικαλυπτόμενων) συνασπισμών. Έτσι, αρχικά προτείνουμε ένα καινοφανές σχήμα συνοπτικής αναπαράστασης, το οποίο ονομάζουμε "Relational Rules", και το οποίο επεκτείνει την εξαιρετικά γνωστή αναπαράσταση "MC-nets" σε συνεργατικά παίγνια με επικαλυπτόμενους συνασπισμούς. Στη συνέχεια παρουσιάζουμε μία νέα μέθοδο λήψης αποφάσεων για αποκεντροποιημένο σχηματισμό επικαλυπτόμενων συνασπισμών, η οποία χρησιμοποιεί, για πρώτη φορά στη βιβλιογραφία της δημιουργίας συνασπισμών, μια επιτυχημένη προσέγγιση μή επιβλεπόμενης μηχανικής μάθησης "Probabilistic Topic Modeling" (¨Πιθανοτικής Θεματικής Μοντελοποίησης¨). Τα πειράματά μας δεικνύουν ότι οι πράκτορες, ερμηνεύοντας τους συνασπισμούς ως έγγραφα, μπορούν αποτελεσματικά να μάθουν επικερδή πρότυπα συνεργασίας (ή "topics") με αποδοτικό τρόπο.

Acknowledgements

First, I would like to thank my advisor, Georgios Chalkiadakis, for all his support and the disciplines he has taught me. He has definitely had a fundamental effect on my decision to pursue further graduate studies. His giving me the opportunity to take my first steps in research has been invaluable.

Moreover, I want to thank Michail Lagoudakis and Ioannis Refanidis for their support and for being great teachers. I am also going to miss the insightful conversations I had with Vasilis Samoladas. I believe that it is imperative to not forgetting where our roots lie, and thus I would like to thank my undergraduate thesis advisor, Ilias Sakellariou. At this point, I would like to pay my tribute to those who make knowledge accessible to the inexperienced ones. I do not think that I would have ended up here if it weren't for the textbooks of Michael Wooldridge in the first place.

Finally, I cannot thank my parents and sister enough; Kostas, Fragka and Foteini have always been there for me with all their love.

"The sleep of reason produces monsters."

Francisco Goya

Contents

List of Figures

List of Tables

List of Algorithms

To those who consciously go against all odds.

Chapter 1

Introduction

Multi-agent systems (Wooldridge, 2009) has risen as a multidisciplinary field whose roots lie in artificial intelligence, but spread further beyond to economics, sociology and distributed systems. In particular, early research focused on formal methods and logic-based approaches. A number of models for building cognitive agents was initially proposed, with Beliefs-Desires-Intentions (BDI) (Rao and Georgeff, 1991; Cohen and Levesque, 1990) emerging as the most promising one. However, despite that aspects of major consideration in multi-agent systems, such as bounded rationality (Bratman, Israel, and Pollack, 1988), had already been recognized, research focus soon took a turn to aspects of computational economics (Boutilier, Shoham, and Wellman, 1997). At that time, it was evident that a formal theory for the description of agent interactions was needed.

Game theory (Osborne and Rubinstein, 1994; Myerson, 2013) not only appeared as the most prominent candidate for formalizing economic agents, but its adoption gave birth to additional research questions and directions in multi-agent systems, as well. This was not accidental, since the interplay between computer science and game theory had already been transparent, leading to the emergence of algorithmic game theory (Nisan et al., 2007). Therefore, it did not take long for cooperative game theory (Chalkiadakis, Elkind, and Wooldridge, 2011) to attract the interest of a plethora of researchers. One of the main points of focus of cooperative game theory is the study of stability, i.e., the conditions under which no subset of agents has an incentive to deviate from a given coalition structure. Naturally, this gives rise to the question of how the agents proceed in forming coalitions in the first place, which is termed as the coalition formation problem. The coalition formation problem is especially concerned with finding the best possible way, in terms of gained utility, for agents to form teams (coalitions) and proceed to joint actions. Both centralized and decentralized approaches have been studied in multi-agent systems literature. The difference between these two is essentially in that in the former approach a solution is given *for* the agents, while in the latter the agents have to make decisions *by themselves*.

In this thesis, we study agent decision-making methods for coalition formation under uncertainty, focusing on environments where agents can participate in more than one coalition simultaneously. As such, the approach that this thesis adopts is a decentralized one. There is a number of reasons that this approach is preferred to a centralized one. First, providing a solution to an agent disturbs the notion of autonomy, which is of major importance in the definition of the characteristics of an agent (Wooldridge and Jennings, 1995). Secondly, it is highly unlikely that even if all agents of a system trust an "one-above-all entity", that entity will manage to inform everyone regarding the right decision on time. This becomes apparent once one thinks of a future where agents all over the Internet need to take decisions during ongoing negotiations or have to rapidly form beneficial coalitions in electronic marketplaces (Li et al., 2005). Furthermore, we are interested in environments where an agent can participate in multiple coalitions at the same time. This can be the case when agents hold some amount of divisible resource(s), e.g., money, time, computational power. Therefore, the emergence of the concept of *overlapping* coalitions (Chalkiadakis et al., 2008; Chalkiadakis et al., 2010) raises a number of interesting questions.

Moreover, learning in environments with high degree of uncertainty has been a major task in artificial intelligence and machine learning (Bishop, 2006). Naturally, research on multi-agent systems could not have been unaffected. It is a major challenge to keep on improving the way that an agent learns how to act, in the presence of other agents (Tuyls and Weiss, 2012).

The insight on the reasons that led to the conduction of this thesis is provided in the next subsection. Finally, at the end of the chapter, the contributions of this thesis are briefly overviewed, followed by its outline.

1.1 Motivation

Cooperative transferable utility (*TU*) games have been widely studied (Chalkiadakis, Elkind, and Wooldridge, 2011) as they provide a rich framework for cooperation and coordination among rational utility-maximizing agents. They model a plethora of real-world scenarios where agents need to jointly tackle tasks that result to some utility (*coalitional value*) awarded to the team (*coalition*). Their applications range from efficient buyer coalitions in e-marketplaces (Yamamoto and Sycara, 2001) to smart grid technologies (Chalkiadakis et al., 2011; Robu et al., 2012) and the power distribution in council bodies (Elkind, Pasechnik, and Zick, 2013). The importance of their study is also indicated by the fact that in many real-world settings it may be impossible for agents to achieve their goals on their own, or that it is simply more profitable to form coalitions with other agents.

Even though the vast majority of literature assumes that coalitions have to be disjoint—i.e., an agent can belong in only one coalition—this assumption does not hold in many realistic scenarios, in which an agent can be in several coalitions simultaneously (Chalkiadakis et al., 2010). Therefore, research on *overlapping* coalition formation (*OCF*) has been greatly neglected. However, it is very common for agents that operate in real-world environments to hold some divisible resource. The introduction of *coalition structures with overlapping coalitions* creates a whole new research direction in cooperative game theory and coalition formation. This alone is exciting, not to mention the impact that the relevant results can bring in the application of multi-agent systems in the real world.

Naturally, uncertainty is inherited in most realistic scenarios. We believe that neglecting it is a major pitfall. Usually, agents have incomplete information regarding either the resource investment of the other agents, that is needed for the completion of coalitional tasks, or the relations that agents have and which determine the coalitional values. Therefore, this thesis is concerned with providing effective methods for *overlapping coalition formation under uncertainty*. In particular, we approach this problem by exploiting concepts from two distinct fields: Probabilistic Computing (and in particular Probability Inequalities) and Probabilistic Topic Modeling.

Randomized algorithms (Motwani and Raghavan, 1995; Mitzenmacher and Upfal, 2005) lie at the heart of the foundations of computer science. Their popularity stems from the beneficial trade-offs that they provide towards attacking hard problems. Probabilistic computing, in general, offers a plethora of valuable concepts and solutions. Motivated by this fact, in this thesis we exploit probability inequalities, and provide three methods for overlapping coalition formation that build on obtained probability bounds.

Moreover, interconnected electronic societies (Papadimitriou, 2001) and social networks offer a natural environment for the completion of goals of individuals, most of the times by completing tasks through common actions. Therefore, it is intriguing to conduct research on coalition formation problems where the structure of the network affects the plausibility of cooperation among the agents.

Learning is one of the fundamental problems in artificial intelligence (Russell and Norvig, 1995). The vast majority of approaches in single and multi-agent environments that are modeled as games (cooperative or not), is concerned with Reinforcement Learning algorithms (Sutton and Barto, 1998). However, the field of *machine learning* has, especially in recent years, developed a number of supervised and unsupervised methods which exhibit robust performance in many application domains—notably in domains involving big data (Bishop, 2006; Barber, 2012; Michalski, Carbonell, and Mitchell, 2013; Sebastiani, 2002; Cortes and Vapnik, 1995). In this thesis, we employ an unsupervised machine learning paradigm, Probabilistic Topic Modeling (PTM) (Blei, 2012), motivated by its success on the problems it has been applied to. Therefore, as we want to approach multi-agent learning for overlapping coalition formation under uncertainty, we have found the exploitation of PTM to be both intriguing and effective. It is the first time that PTM is used in a multi-agent problem, and, thus, this alone is a source of motivation.

As mentioned above, overlapping coalition formation has, surprisingly, attracted little of attention in cooperative game theory and multi-agent systems literature. Furthermore, overlapping coalition formation under uncertainty has not been studied at all. Therefore, dealing with a problem that is both interesting and has the potential to offer a number of solutions in real-world applications, for the first time in literature is quite a source to derive motivation from.

1.2 Contributions and Thesis Outline

This thesis essentially consists of two conceptually interlinked works, which both attack the overlapping coalition formation under uncertainty problem, but from different perspectives. Thus, the contributions of each are distinctly presented in this section. However, it has to be emphasized that this thesis consists the very first approach in overlapping coalition formation under uncertainty.

The major contribution of the first work consists of three novel methods which benefit from obtained probability bounds for assessing the ability of teams of agents to accomplish coalitional tasks. It is the first time that probability bounds are computed for coalition formation under uncertainty. Furthermore, we provide a way for modeling agent beliefs regarding the contribution of other agents to potential coalitions, when they do not know with certainty the resource quantity the others possess. Moreover, we prove an improvement of the Paley-Zygmund inequality (Paley and Zygmund, 1932), a probability inequality that our first method exploits and that is used for deriving lower probability bounds. Despite that this improvement is reported at some places on the Web, we did not manage to find a reference to a formal proof, and, thus, we had to prove it ourselves. Finally, the contributions of this work include an overlapping coalition formation protocol, with the motivation behind its design stemming from real-world scenarios.

In the second work, we exploit probabilistic topic modeling for overlapping coalition formation under uncertainty. This work essentially presents an entirely novel agent-learning paradigm for coalition formation. We consider the case where there is structural uncertainty, and thus agents do not know how well they can cooperate with others. We show how an agent can gain and exploit knowledge by providing a correspondence between probabilistic topic modeling and learning the profitability of the coalitions. We believe that it can also be employed in non-cooperative environments (though we did not test it in such), and thus serve as a widely accepted technique in multi-agent learning. A key point on the effectiveness of probabilistic topic modeling for overlapping coalition formation is the way via which agents interpret coalitions as documents, and so they can proceed with their learning process. Furthermore, in order to concisely represent the (unknown) underlying inter-agent collaboration structure, we have extended the much celebrated Marginal Contribution networks (MC-nets) (Ieong and Shoham, 2005) representation scheme to cooperative games with overlapping coalitions.

The rest of the thesis is structured as follows:

- In Section 2 we provide basic background on coalition formation, including solution concepts in cooperative game theory, concise representation schemes for cooperative games, and work on centralized coalition formation (coalition structure generation). Moreover, we overview the literature on overlapping coalition formation, uncertainty and learning in, mainly, cooperative games, and stability results on cooperative games on graphs and networks.

- In Section 3 we present the main results from literature on probability inequalities. That section thus provides basic background before the presentation of the main part of our work on probability bounds for overlapping coalition formation.

- In Section 4 we provide the basic background on probabilistic topic modeling. In particular, we introduce the terminology that is widely used in the document analysis domain, and describe the Latent Dirichlet Allocation and the online Latent Dirichlet Allocation algorithms used in our work.

- In Section 5 we present the first of our two major contributions. It includes our three proposed methods for obtaining and deriving probability bounds for overlapping coalition formation— namely, $IPZY$, $CH2S$, and HF—and an overlapping coalition formation protocol through which agents can iteratively form (overlapping) coalitions. It concludes with the results of the experiments we conducted on both a random graph of 300 nodes (agents) and a large-scale real social network from Facebook of 4039 users.

- In Section 6 we present our second major contribution, which is concerned with overlapping coalition formation via probabilistic topic modeling. After providing the basic background on PTM and Online Latent Dirichlet Allocation (the PTM method we exploit), we present *Relational Rules*, which consist the extension of MC-nets to cooperative games with overlapping coalitions. Next, we show how agents can learn by interpreting coalition as documents and a present `OVERPRO`, a method for taking decisions based on the the gained knowledge. We also present a Q-learning algorithm that we have developed and which we use as a baseline. Our experiments show that `OVERPRO` decisively outperforms the baseline and is highly effective regarding the social welfare and the efficient investment of agent resources.

- Finally, Section 7 summarizes the results of this thesis and outlines future work.

Chapter 2

Coalition Formation and Cooperative Games

Coalition formation is the problem where the self-interested agents of a system have to form effective teams (coalitions) in order to achieve their goals. Effectiveness is expressed in terms of gained utility which agents seek to maximize. Therefore, in a typical coalition formation environment, autonomous agents act in a decentralized manner seeking to gain as much utility as possible. The importance of the coalition formation problem rises from the fact that in a plethora of real-world environments agents cannot satisfy their goals on their own, or they simply can gain more by cooperating with other agents. Coalition formation is being formally treated within cooperative game theory (Myerson, 2013; Chalkiadakis, Elkind, and Wooldridge, 2011); and much of the work in the field is concerned with the development and study of solution concepts related to *stability*, such as the *core* (Gillies, 1959). Despite that the approach of this thesis is not game-theoretical per se, in this chapter we overview the basic concepts of cooperative game theory, since they are closely related to our work. Furthermore, the basic principles of multi-agent systems (Wooldridge, 2009) indicate that agents have to act autonomously and in a decentralized manner. However, in order to prevent the occurrence of any misunderstandings, it has to be emphasized that *norms* and *rules* (Dechesne et al., 2013) do exist in multi-agent systems. For instance, auctions (Shoham and Leyton-Brown, 2008) consist a broadly accepted decentralized mechanism for item allocation, but agents that participate in those have to take actions in strictly specified time moments.

Despite that the coalition formation problem is specified for autonomous agents in distributed environments, works in literature often mistakenly refer to the problem of finding the coalition structure that maximizes the social welfare (sum of agents' utility), via a centralized algorithm or heuristic, by the same name. The proper way that one should refer to that problem is *coalition structure generation* (Rahwan et al., 2015), and it often proves to be valuable, as it can serve as a quality indicator of a decentralized solution, in the spirit of price of anarchy (Koutsoupias and Papadimitriou, 1999) and price of democracy (Chalkiadakis et al., 2009). It is, thus, also overviewed in this chapter.

2.1 Transferable Utility Games

Game theory offers a mathematical framework for the analysis of the interaction of self-interested agents. Therefore, a game, despite that it may appear as an abstract notion at first, has a solid definition. Cooperative game theory is concerned with both *transferable* and *non-transferable* utility games (Chalkiadakis, Elkind, and Wooldridge, 2011) (though technically this equation of terms is not entirely correct, due to the existence of partition function games). This thesis approaches only the former, which are often referred to as *characteristic function games*. In non-transferable utility games, a coalition of agents cannot arrange the distribution of the payments. Instead, the gain of a participant is defined based solely on her preferences (Chalkiadakis, Elkind, and Wooldridge, 2011). Hedonic games (Brandt et al., 2016; Aziz and Brandl, 2012; Elkind and Wooldridge, 2009), for instance, consist a class of non-transferable utility games where the utility that an agent gains by his participation in a coalition is defined by the identity of the other participants. In a sense, the company of the members of a group offers some amount of "pleasure" to a participant (hedony stand for joy, or pleasure, in greek).

A transferable utility game $G = (N, u)$ is defined by a set of agents $N = \{1, \ldots, n\}$ and a characteristic function $u : 2^N \to \mathbb{R}$. Therefore, $u(C)$, where $C \subseteq N$, is the value earned by coalition C. The *allocation vector* is denoted as $x = (x_1, \ldots, x_n)$, and defines the *payoff distribution* among the agents of the game. Often, $x(C)$ is used as a notation for the allocation of the utility earned by the members of C. A coalition structure CS is an exhaustive partition of N, and, thus, consists of a set of coalitions $\{C_1, \ldots, C_k\}$, such that $C_a \cap C_b = \emptyset$, for any two coalitions $C_a, C_b \subseteq CS$, and $\cup_{C \in CS} = N$. Naturally, an allocation x is feasible if $\sum_{i \in C} x_i \leq u(C)$, for every $C \in CS$. A coalition that consists of a single agent is called a *singleton*. Agent i can earn his reservation value rv_i by forming a singleton. Therefore, a coalition structure CS consists of at least one coalition, the grand coalition ($C = N$, where all agents join forces together), and at most n coalitions, where each agent forms a singleton. It is apparent that characteristic function games are of great interest since agents are able to make binding agreements on the distribution of the utility, which are usually conducted via a specified interaction protocol.

Some subclasses of transferable utility games have played an important role on cooperative game theory research, as their properties have both theoretical and practical implications. In *superadditive games* it is always profitable for two coalitions to merge, i.e., $u(C \cup D) \geq u(C) + u(D), C, D \subseteq N$. Therefore, in such games it is natural to expect the grand coalition to form. The key question in superadditive games is on how the utility $u(N)$ of the grand coalition should be allocated to the agents with respect to a stability solution concept (an overview follows on the next section). Another subclass that has played a significant role, especially in applications that involve elections and political parties, is that of *simple games*. In simple games a coalition can be (strictly) either winning or losing, i.e., $u(C) = 1$ or $u(C) = 0$. A player i is said to be *dummy* if his membership in a coalition makes no difference in the gained utility, i.e., $u(C \cup i) = u(C)$, for every $C \subseteq N \setminus i$. Player i is veto if $u(C) = 0$ for every $C \subseteq N \setminus i$, i.e., a coalition cannot win if i is not a member of it. However, this does not imply that every coalition that includes i has to be winning. Furthermore, there might zero to n veto players. Simple games are related to *weighted voting games*, where agents have weights and coalition C wins, i.e., $u(C) = 1$, if the sum of the weights of its members surpasses a certain quota (threshold).

2.2 Solution Concepts

Stability is an issue of major concern when self-interested autonomous agents seek to achieve their goals. Therefore, various concepts of stability have been developed since the very early years of game theory. In non-cooperative game theory Nash-equilibrium (Nash, 1951) is probably the most well-celebrated concept of stability, with significant results on its approximability having been recently found (Daskalakis and Papadimitriou, 2015).

Stability in cooperative game theory takes under consideration *coalitions* of agents, rather than individual agents—a single agent is represented as a singleton. A *coalitional configuration* is characterized by a coalition structure CS and a vector of payoffs x. One of the most well-known solution concepts is the *core* (Gillies, 1959), and it consists the closest analogy to Nash-equilibrium, in terms of non-cooperative game theory. An outcome (CS, x) is in the core if no coalition can benefit by deviating. Therefore, the core is set that includes all the outcomes (CS, x) such that $x(C) \geq u(C)$ for every $C \subseteq N$. The core might seem as an ideal concept of stability in cooperative games, but it has one major drawback: in the general case, it is computationally intractable (Rapoport, 1970; Conitzer and Sandholm, 2003).

The kernel (Davis and Maschler, 1965) is another solution concept that is related to the *surplus* among every pair of agents. One of the attractive properties of the kernel is that it admits good polynomial-time approximation schemes (Shehory and Kraus, 1999; Bistaffa et al., 2015). Other solutions concepts are the *nucleolus* (Schmeidler, 1969), which is based on the notion of deficit, the *bargaining set* (Aumann and Maschler, 1964), which is based on objections and counterobjections, and the *stable set* (Von Neumann and Morgenstern, 1947), which is based on the notion of dominance. In contrast to the ones already presented, the *Shapley value* (Shapley, 1952) is a value-based solution concept, but its exact computation cannot be efficiently performed.

Solution concepts are fundamental in cooperative game theory, but the computation of the most is intractable. In particular, the core, which is the most attractive solution concept, not only cannot be efficiently computed, but it might even be empty. Given this, in this thesis we focus on the problem of decision-making that an agent has to tackle when there is uncertainty, without proceeding in the computation of solution concepts.

2.3 Consice Representation Schemes

It is apparent that the utility function u of a transferable utility game $G = (N, u)$ plays a central role on the decisions of the agents. However, since u is defined as $u : 2^N \to \mathbb{R}$, making an optimal decision is, in general, intractable, as it involves $O(2^n)$ computation steps. Therefore, the need for succinct representation formalism has led to the development of *concise representation schemes* (Chalkiadakis, Elkind, and Wooldridge, 2011). Such schemes are divided in two main categories, (i) schemes that never fail to compactly represent a game, but cannot represent every possible game, and (ii) schemes that can represent any game but can fail to do so in a compact way.

Combinatorial optimization games always provide succinct representation, but cannot model every game, and are largely based on graph theory (Myerson, 1977). *Induced subgraph games* (Deng and Papadimitriou, 1994) consist the first such representation. In this one, a game is described by a weighted undirected graph $G = (N, E)$, where each node $i \in N$ represents an agent. The value $u(C)$ of coalition C is defined as $u(C) = \sum_{\{i,j\} \in C \cap E} w_{i,j}$, where $w_{i,j}$ is the weight of the edge $\{i, j\}$. Despite that computing the value of a coalition is tractable (there is a polynomial time and space algorithm), induced subgraph games are not *complete*, since the value of a coalition is often not defined by pairs of agents only. However, induced subgraph games are especially valuable in modeling agent interactions in social networks. Other subclasses of combinatorial optimization games are the *network flow games* (Kalai and Zemel, 1982a; Kalai and Zemel, 1982b), the *matching games* (Deng, Ibaraki, and Nagamochi, 1999), the *minimum cost spanning games* (Bird, 1976), and the *facility location games* (Goemans and Skutella, 2004).

Complete representation schemes owe their characterization to the fact that they can represent every transferable utility game However, they do not always succeed in achieving succinctness. One of the most popular is the *Marginal Contribution networks* (MC-nets) (Ieong and Shoham, 2005). In MC-nets, coalitional games are represented by a set of rules of the form:

$$Pattern \to value$$

where $Pattern$ is a conjunction of literals (representing the participation or absence of agents), and applies to coalition C if it satisfies $Pattern$, with $value \in \mathbb{R}$ being added to the coalitional value of C. For instance, if the set of the rules of a game consists of $\{1 \wedge 2 \wedge 3\} \to 8$, $\{1 \wedge 2\} \to -6$ and $\{2\} \to 4$, then the value of coalition $C = \{1, 2, 3\}$ is $u_C = 8 - 6 + 4 = 6$, since all rules apply, and the value of coalition $C' = \{1, 2\}$ is $u_{C'} = -6 + 4 = -2$, since only two of the rules apply. Therefore, the basic idea of this scheme is that the relations among the agents can be described by the rules mentioned above. The succinctness they offer strongly depends on the number of rules that exist in a cooperative game. In many settings, it is natural to expect that not all of the subsets of agents are characterized by a rule. Nevertheless, this might not hold, and, thus, the number of rules in MC-nets can be exponential in the number of the agents. Another interesting concise representation scheme that is based on *synergies* among agents is presented in (Conitzer and Sandholm, 2006). That scheme relies on super-additivity, and allows efficient checking of whether a given outcome is in the core of the game. Other representation schemes rely on agent types. In (Ueda et al., 2011) the presented scheme exploits *recognizable equivalence* of the agents, where agent i and j are recognizably equivalent if for any coalition C, such that $i, j \notin C$, it holds that $u(C \cup \{i\}) = u(C \cup \{j\})$. Therefore, letting $T = \{1, \ldots, t\}$ be the set of recognizable types, coalition C is characterized by the vector $n_C = \langle n_C^1, \ldots, n_C^t \rangle$, where n_C^k is the number of agents of type k that are in C. Other complete concise representation schemes are based on skills of agents (Ohta et al., 2006) and algebraic decision diagrams (Bahar et al., 1997; Aadithya, Michalak, and Jennings, 2011).

2.4 Coalition Structure Generation

As mentioned in the introduction, it is valuable to consider settings where agents do not take action according to their free will, but follow the instructions of a central designer who is interested in maximizing the social welfare, i.e, the total utility earned by the agents. Despite that taking this as a norm of agent decision making would disturb the notion of autonomy, which lies in the foundations of multi-agent systems (Wooldridge and Jennings, 1995), studying *coalition structure generation*, i.e. centralized coalition formation, is important in evaluating the quality of the solutions given by decentralized approaches and equilibria (Koutsoupias and Papadimitriou, 1999; Chalkiadakis et al., 2009).

One of the early works in coalition structure generation presented a search algorithm that establishes lower bounds on the quality of the solution (Sandholm et al., 1999). Other approaches are based on dynamic programming (Rahwan and Jennings, 2008) and integer-partition (Rahwan et al., 2009). In particular, dynamic programming enables the reduction of number of partitions from $O(n^n)$ to $O(3^n)$. Subspace search and integer-partition techniques that build on dynamic programming have shown to perform even better.

Graph representation is useful when there are restrictions on the coalitions that can be formed, since in real-world applications one agent might be unable to directly reach another due to constraints on communication. Coalition structure generation on graphs was studied in (Voice, Polukarov, and Jennings, 2012), providing results that rely heavily on tree decomposition techniques. This approach is actually quite common in games with an underlying feasibility-formation structure, where the notion of *treewidth* is usually exploited. Intuitively, treewidth is a measure of telling how far a graph is from being a tree. Many problems that are hard in the general case become easy when the treewidth is small (Bodlaender and Koster, 2008). A recent work on coalition structure generation on graphs (Bistaffa et al., 2014) exploits the concept of *edge contraction* which allows the efficient representation of the search space. That proposed algorithm is *anytime*, i.e., it can be stopped during runtime and still provide a solution, but as long as it keeps on running it can only improve it. Furthermore, it is parallelizable, and it is shown to experimentally perform better than the previous state-of-the-art algorithm (Voice, Ramchurn, and Jennings, 2012). Coalition structure generation has many real-world applications, varying from smart grid (Oliveira Ramos and Bazzan, 2012) to ride sharing (Bistaffa et al., 2015) problems. There is a (relatively) recent survey on coalition structure generation (Rahwan et al., 2015).

2.5 Overlapping Coalition Formation

Overlapping coalition formation was initially studied in (Shehory and Kraus, 1998), which provided an approximate solution to the corresponding optimal coalition structure generation problem (Rahwan et al., 2015), in a setting where the costs of the coalitions and the capabilities of the agents are globally announced. The proposed method of (Shehory and Kraus, 1998) employs concepts from combinatorics and approximation algorithms. Though related, our approach differs in that it is *decentralized*, since the (overlapping) coalitions are formed by the agents themselves, and are *not* provided for the agents by an algorithm. The subsequent work of (Dang et al., 2006) presented an application of overlapping coalitions in sensor networks. An approximate greedy algorithm with worst-case guarantees is introduced, and constitutes a real-world example of employing overlapping coalitions. However, in that work, the agents do not form coalitions acting in a completely autonomous manner, since they are hardwired to agree on taking a specific action (regarding the choice of an agent), in a step of the algorithm.

As illustrated by the work of (Chalkiadakis et al., 2008; Chalkiadakis et al., 2010), in overlapping coalition formation games an agent can be part of a number of coalitions simultaneously, and thus coalition structures are much more complex than in games without overlapping coalitions. One reason that gives rise to this complexity is the fact that the number of different coalition structures cannot be enumerated; even a single agent can form an infinite number of singletons. Furthermore, the concept of stability in games with overlapping coalitions is quite different, since it is not just the membership of an agent in a coalition that matters, but the degree by which she participates in that. Therefore, a payoff-allocation has to be determined by taking under consideration much more complex structures (Chalkiadakis et al., 2010). Additionally, in such settings, it is not just agents that can deviate, but whole coalition structures, since coalitions of agents can withdraw just a portion of their resource from the formed coalitions. Such complex structures give rise to more challenging problems than those imposed by non-overlapping coalition formation games (Chalkiadakis et al., 2010).

Indeed, the formal definition of cooperative games with overlapping coalitions was not until (Chalkiadakis et al., 2008; Chalkiadakis et al., 2010). In that model, an agent is allowed to be part of more than one coalition, by contributing some portion of his resources to each one; and thus a partial coalition is given by a vector $r^C = (r_1^C, \ldots, r_n^C)$, where $r_i^C \in [0,1]$ denotes the resource fraction which agent i contributes to coalition C, so $r_i^C = 0$ means that agent i is not part of coalition C. In this way, an *overlapping coalition formation game* (OCF-game) is given by (i) a set of players $N = \{1, \ldots, n\}$ and (ii) a characteristic function $v : [0,1]^n \mapsto \mathbb{R}$. One point of focus of the works mentioned above, is the characterization of the core for OCF games. Especially, the notion of *balancedness* (Bondareva, 1963; Shapley, 1967), in the context of overlapping coalition formation, is studied. Balancedness consists a condition that an outcome (CS, x) has to satisfy, in order to belong to the core. In (Chalkiadakis et al., 2010) it is shown that the condition of balancedness is more complicated in OCF settings, as the linear program that describes core allocations requires a larger set of constraints.

In (Chalkiadakis et al., 2010), an expressive class of OCF games, *threshold task games* (TTGs), is also put forward. A threshold task game $G = (N; w; t)$ is given by (i) a set of agents $N = \{1, \ldots, n\}$, (ii) a vector $w = (w_1, \ldots, w_n) \in \mathbb{R}^+$ denoting the quantity of resources the agents possess and (iii) a list $t = (t_1, \ldots, t_l)$ of task types where each t_h is described by a threshold value $T_h \geq 0$ and a utility $v_h \geq 0$, so a task type is denoted $t_h = (T_h, v_h)$. In TTGs, agents form coalitions to complete tasks, in order to gain utility v_h by fulfilling the requirement T_h for t_h.

In a series of works (Zick and Elkind, 2011; Zick, Chalkiadakis, and Elkind, 2012; Zick, Markakis, and Elkind, 2014) stability with respect to the behaviour of non-deviating players towards deviators is studied. Various variations of the core are developed, namely the *conservative core*, the *refined core*, and the *optimistic core*, and the approach is based on the notion of *arbitration functions*, which define the payoff of the deviators according to the attitude of the non-deviators. The circumstances under which these solution concepts are non-empty are also studied.

Furthermore, the most related class of games to the cooperative games with overlapping coalitions is the one of *fuzzy coalitional games* (Aubin, 1981). In a fuzzy game, an agent can be part of a coalition at various levels. Thus, the coalitional value of $C \subseteq N$ is defined by the level at which the agents have joined C. While there is a number of differences between overlapping coalition formation and fuzzy games, the greatest one is that in fuzzy games the core is the only acceptable outcome. Finally, coalition structure generation with overlapping coalitions is studied in (Zhang et al., 2010), where a metaheuristic is developed, based on particle swarm optimization (Eberhart, Kennedy, et al., 1995).

2.6 Uncertainty and Learning

Stochasticity in the value of payoffs in non-overlapping cooperative games has been studied in (Suijs et al., 1999), in a setting where agents have different preferences over a set of random variables. The focus of this study is on core-stability. Bayesian coalitional games are introduced in (Ieong and Shoham, 2008), where suitable variations of the core are also defined. In (Kraus, Shehory, and Taase, 2003; Kraus, Shehory, and Taase, 2004) agents have incomplete information regarding the costs that the other agents incur by performing a task within a coalition, while the formation of the coalitions takes place through information-revealing negotiations and the conduction of auctions. Nevertheless, the formation of overlapping coalitions is not allowed in (Kraus, Shehory, and Taase, 2003; Kraus, Shehory, and Taase, 2004). An approach to coalition formation via bargaining where an agent has beliefs regarding the types of the others is presented in (Chalkiadakis and Boutilier, 2007). It is shown that the game can be described as a Bayesian game in extensive form, and a heuristic for approximating the optimal solution is given. A link between non-cooperative and cooperative solution concepts in coalitional bargaining under uncertainty is studied in (Chalkiadakis and Boutilier, 2007).

One of the very early attempts in approaching learning in cooperative settings was presented in (Claus and Boutilier, 1998), where the dynamics of a set of reinforcement learning (RL) algorithms (Sutton and Barto, 1998) were studied. In (Kapetanakis and Kudenko, 2002) a heuristic for action selection strategy in Q-learning in cooperative environments was developed, which is experimentally shown to almost always converge to a desirable state. A Bayesian approach to reinforcement learning for coalition formation is presented in (Chalkiadakis and Boutilier, 2004), along with the introduction of a variation of the core. The recent work of (Balcan, Procaccia, and Zick, 2015) explores a PAC (probably approximately correct) model for obtaining theoretical predictions for the value of coalitions that have not been observed in the past. Common classes of cooperative games are examined there regarding their PAC learnability, among which *Threshold Task Games* (TTGs) (Chalkiadakis et al., 2010), a class of OCF games already presented. The links between evolutionary game theory and multi-agent reinforcement learning is the topic of study in (Tuyls and Parsons, 2007; Kaisers and Tuyls, 2009). For a discussion on the fundamental connections between multi-agent learning and game theory, and the impact that the latter has on the former, one should study (Shoham, Powers, and Grenager, 2007) and (Stone, 2007).

Multi-agent learning in non-cooperative games (Fudenberg and Levine, 1998) has been studied for a longer time. Much of the early seminal work (Littman, 1994; Hu and Wellman, 1998; Hu and Wellman, 2003) is interested in the study of Q-learning algorithms and their convergence to Nash equilibria (Nash, 1951). In particular, the algorithm presented in (Hu and Wellman, 1998) is shown to converge to a Nash equilibrium if every state and action has been visited infinitely often and the learning rate satisfies some conditions regarding the values it takes over time. Therefore, despite that in theory convergence can be guaranteed, in practice the first assumption cannot hold. Overall, the literature on multi-agent learning (Tuyls and Weiss, 2012), in both cooperative and non-cooperative settings, is largely concerned with the study of reinforcement learning algorithms.

2.7 Stability in Networks

Cooperative games on graphs were introduced in the seminal paper of (Myerson, 1977). In that work the notion of connectedness of the agents was presented, relying on the structure of the underlying graph, whose nodes represent the agents of the game and a coalition is feasible only if the induced subgraph is connected. A plethora of complexity-theoretic results on transferable utility games with graph restrictions are presented in (Chalkiadakis, Greco, and Markakis, 2016). The stability of a game with respect to the type (line, tree, cycle) of the underlying graph structure was studied in (Chalkiadakis, Markakis, and Jennings, 2012). In that work, partition function games (i.e., games where the value of a coalition depends on the other formed coalitions) are also taken under consideration, while a Bayesian extension of those is proposed, too. Stability in hedonic games where the communication among the agents is restricted by the graph structure was recently studied in (Igarashi and Elkind, 2016). The positive results of that work depend on graph acyclicity, by large. Most results, just as in characteristic function games, are negative, since most problems are NP-$Complete$. Non-cooperative solution concepts for coordination games on graphs, which are closely related to cooperative games, were studied in (Apt, Simon, and Wojtczak, 2015). The results are related to the existence of Nash equilibria, and it is shown that the underlying problems are NP-$Complete$. Now, stability on overlapping coalition formation games on graphs has been studied in (Zick, Chalkiadakis, and Elkind, 2012). The derived positive results hold only under very limiting restrictions, and are related to the maximum number of agents that can participate in a coalition and the treewidth of the graph.

Chapter 3

Probability Inequalities

Given that digital computers are completely deterministic machines, the addition of randomness in computation seems quite frustrating. At first sight, it may not be clear why one would desire to get an outcome that comes out of a distribution, rather than a deterministic one. Still, randomness, and thus probability theory, has played a major role in computer science. Computational complexity characterizes the difficulty of problems based on every possible instance. Therefore, it might be the case that only a fraction of instances of a problem are hard to solve. In that case, it sounds a like good idea to get an *approximate* solution to that problem *most* times. Indeed, worst-case complexity might fail to capture the underlying difficulty of a problem, while average-case complexity (Impagliazzo, 1995) can be more suitable. Developing a *randomized algorithm* (Motwani and Raghavan, 1995; Mitzenmacher and Upfal, 2005), one where the toss of a coin is involved, has become quite a popular approach, closely related to that of *approximation algorithms* (Vazirani, 2001). A *Las Vegas* randomized algorithm is one that always returns correct results, while a *Monte Carlo* randomized algorithm is one whose output may not be correct, with some probability which is typically small. The latter is most commonly used in practice. Thus, the fundamental idea behind randomized algorithms is that a polynomial algorithm can be developed for an *intractable* problem, by taking a decision at random at a step, with a (known) probability of returning a wrong result. Sometimes, the quality of the solution can be bounded.

A major concept in randomized algorithms and decision making under uncertainty is that of *probability inequalities*. Probability inequalities provide bounds on the values of probabilities of events that are related to tails of distributions. They are commonly used in the analysis of algorithms, as they are in many cases a useful tool for bounding time complexity. However, their applications are not limited to that. In particular, in the context of this thesis, we exploit probability inequalities for agent decision-making, as they provide a means of approximating the values of probabilities that are otherwise expensive to compute.

3.1 Expectation and Variance of Random Variables

A random variable is a function whose domain is the set of events of the sample space, and takes values from the set of (greater or equal to zero) real numbers. For instance, the domain of a variable that represents the toss of a (fair) six-sided dice is $\Omega = \{1, 2, 3, 4, 5, 6\}$, and for an event $x \in \Omega$, $P(A = a) = 1/6$. Formally:

Definition 1. *A random variable (r.v.) A is a real-valued function on sample space Ω, i.e., $A : \Omega \to \mathbb{R}$. A random variable can be either discrete or continuous, that is it can take values from either a discrete space (or a countably infinite one) or a continuous space.*

Furthermore, independence of random variables is one of the most fundamental concepts in probability theory. Intuitively, two random variables are independent if the value that the one takes does not affect the one of the other. Formally:

Definition 2. *Two random variables A and B are independent if and only if:*

$$P\big((A = a) \cap (B = b)\big) = P(A = a) \cdot P(B = b)$$

It further holds for independent random variables A and B that:

$$P(A = a | B = b) = P(A = a)$$

We now define *expectation* of a random variable, which is essentially the weighted average of the values of its sample space, where the weight that is given to each value is equal to the probability that the random variable takes that value.

Definition 3. *The expectation of the discrete random variable A, whose sample space is Ω, is:*

$$E[A] = \sum_{a \in \Omega} a \cdot P(A = a)$$

An important property that is related to the sum of a finite number of random variables is the *linearity of expectations*, which is defined as follows:

Definition 4. *The expected value of the sum of the (either dependent or independent) discrete random variables $A_1, \ldots, A_n$ is:*

$$E[A] = E\big[\sum_{i=1}^{n} A_i\big] = \sum_{i=1}^{n} E[A_i]$$

Now, the *variance* of a random variable is a measure that, intuitively, tells how far the random variable is likely to be from its expected value. The *standard* *deviation* of a random variable is the square root of its variance.

Definition 5. *The variance of random variable A is:*

$$Var[A] = E[(A - E[A])^2] = E[A^2] - E[A]^2$$

Therefore, the standard deviation of A is $\sigma[A] = \sqrt{Var[A]}$.

The joint variability of two random variables is termed as *covariance* and is defined as follows:

Definition 6. *The covariance of two random variables A and B is:*

$$Cov[A, B] = E[(A - E[A]) \cdot (B - E[B])]$$

As we have already presented, the expected value of the sum of random variables is simply the sum of the expected value of each distinct random variable, and it is not different for dependent and independent random variables. However, the variance of the sum of two random variables involves one extra term, that of covariance.

Definition 7. *The variance of the sum of two random variables A and B is:*

$$Var[A + B] = Var[A] + Var[B] + 2 \cdot Cov[A, B]$$

For two independent random variables A and B it holds that:

$$Cov[A, B] = 0$$

Therefore, the variance of the sum of two independent r.v.'s A and B is:

$$Var[A + B] = Var[A] + Var[B]$$

3.2 Markov's Inequality

One of the most fundamental probability inequalities is *Markov's inequality*. Despite that the bounds it offers are far from tight (except to the case that only the expected value is known), it often consists the basis for the development of more sophisticated inequalities. Furthermore, it only assumes that the expected value of the random variable, whose tail distribution is computed, is known, and that the sample space consists of events whose values are greater than zero.

Theorem 1. *Let A be a random variable whose sample space includes only non-negative values. Then, for any $\eta \in \mathbb{R}^+$, it holds that:*

$$P(A \geq \eta) \leq \frac{E[A]}{\eta}$$

For a proof of Markov's inequality one can see (Mitzenmacher and Upfal, 2005).

Assume a six-sided dice and let A be the random variable of the outcome of the roll of the dice. Thus, $\Omega = \{1, 2, 3, 4, 5, 6\}$. Since the dice is fair, we know that $E[A] = 3.5$, $P(A \geq 4) = 1/2$, $P(A \geq 5) = 1/3$, and $P(A \geq 6) = 1/6$. According to Markov's inequality $P(A \geq 4) \leq 0.875$, $P(A \geq 5) \leq 0.7$, and $P(A \geq 6) \leq 0.583$. Ideally, we would like the ratio $\frac{bound}{true\_value}$ to get closer to 1, as the value of a increases. However, we see that $\frac{0.875}{1/2} = 1.75$, $\frac{0.7}{1/3} = 2.1$, and $\frac{0.583}{1/6} = 3.5$.

3.3 One-sided Chebyshev's Inequality

An inequality that provides upper bounds on either the lower or the upper tail distribution, with respect to the distance from the expected value is the *one-sided Chebyshev's* inequality. Thus, the one-sided Chebyshev's inequality has two forms and is defined as follows:

Theorem 2. *Let A be a random variable and $\eta \in \mathbb{R}^+$. It holds that:*

$$P(A \geq E[A] + \eta) \leq \frac{Var[A]}{Var[A] + \eta^2}$$

$$P(A \leq E[A] - \eta) \leq \frac{Var[A]}{Var[A] + \eta^2}$$

For a proof on one-sided Chebyshev's inequality one can see (Bast and Weber, 2005) and (Grimmett and Stirzaker, 2001).

In the six-sided dice example, the one sided Chebyshev's inequality yields $P(A \geq 4) \leq 0.853$, $P(A \geq 5) \leq 0.66$, and $P(A \geq 4) \leq 0.538$. Thus, even in this naive problem, we can see that this inequality offers better (tighter) bounds than Markov's inequality.

3.4 Two-sided Chebyshev's Inequality

We can also derive a similar inequality for both tails of the distribution of a random variable known as the *two-sided Chebyshev's* inequality, which is defined as follows:

Theorem 3. *Let A be a random variable and $\eta \in \mathbb{R}^+$. It holds that:*

$$P(|A - E[A]| \geq \eta) \leq \frac{Var[A]}{\eta^2}$$

Proof. We see that $P(|A - E[A]| \geq \eta) = P\big((A - E[A])^2 \geq \eta^2\big)$. Thus, we can apply Markov's inequality, as $(A - E[A])^2 \geq 0$:

$$P((A - E[A])^2 \geq \eta^2) \leq \frac{E[(A - E[A])^2]}{\eta^2} = \frac{E[A^2] - E[A]^2}{\eta^2} = \frac{Var[A]}{\eta^2}$$

□

3.5 Chernoff Bounds

An exceptionally powerful and widely used family of inequalities is the *Chernoff bounds* (Chernoff, 1952). They essentially consist a tool that provides exponentially decreasing bounds on the upper tail distribution. A general guideline in deriving Chernoff bounds is that Markov's inequality can be applied to e^{tA}, where A is a random variable and t is a value that determines the quality of the obtained bound. Thus, a general form Chernoff bound can be derived as follows, where $t > 0$:

$$P(A \geq \eta) = P(e^{tA} \geq e^{t\eta}) \leq \frac{E[e^{tA}]}{e^{t\eta}}$$

However, tighter, and thus better, Chernoff bounds can be derived for some special cases. For instance, for the sum of independent 0-1 random variables, the following bounds for the lower tail hold, where $A_1, \ldots, A_n$ are random variables whose sample space is $\Omega = \{0, 1\}$, $A = \sum_{i=1}^{n} A_i$, $\mu = E[A]$, and $\delta \in (0, 1)$:

$$P(A \leq (1 - \delta)\mu) \leq \left(\frac{e^{-\delta}}{(1 - \delta)^{1-\delta}}\right)^{\mu}$$

$$P(A \leq (1 - \delta)\mu) \leq e^{-\mu\delta^2/2}$$

If $A_1, \ldots, A_n$ are independent random variables such that $P(A_i = 1) = P(A_i = -1) = \frac{1}{2}$, and $A = \sum_{i}^{n} A_i$, then for any $\eta > 0$ it holds that:

$$P(A \geq \eta) \leq e^{-\eta^2/2n}$$

Anther interesting case is when $A_1, \ldots, A_n$ are independent and $P(A_i = 1) = P(A_i = 0) = \frac{1}{2}$. Let $A = \sum_{i=1}^{n} A_i$ and thus $\mu = E[A] = n/2$. The following two Chernoff bounds hold:

For $0 < \eta < \mu$

$$P(A \leq \mu - \eta) \leq e^{-2\eta^2/n}$$

For $0 < \delta < 1$

$$P(A \leq (1-\delta)\mu) \leq e^{-\delta^2 \mu}$$

3.6 Hoeffding's Inequality

A probability inequality for obtaining bounds on both tails around the mean value of the sum of independent random variables was proven in 1963 by Wassily Hoeffding (Hoeffding, 1963). However, it only holds for variables that are bounded by both sides.

Theorem 4. *Let $A_1, \ldots, A_n$ be independent random variables, where all A_i are bounded so that $A_i \in [a_i, b_i]$, and let $A = \sum_{i=1}^{n} A_i$. Then, it holds that:*

$$P(|A - E[A]| \geq k) \leq 2exp\Big(-\frac{2k^2}{\sum_{i=1}^{n}(b_i - a_i)^2}\Big)$$

Naturally, the value of the bound provided by Hoeffding's inequality largely depends on the values of bounds a_i and b_i of the variables. However, it consists one of the most fundamental probabilities inequalities with various applications in statistics (Massart, 2000; Glynn and Ormoneit, 2002).

3.7 Paley-Zygmund Inequality

Most probability inequalities are helpful in providing upper bounds on tail distributions. Deriving inequalities for *lower* bounds is considered harder, in general. Nevertheless, one such inequality is the *Paley-Zygmund* inequality (Paley and Zygmund, 1932) which provides lower bounds on the upper tail of the distribution of a random variable.

Theorem 5. *Let a random variable A (with finite variance), and $0 \leq theta \leq 1$. It holds that:*

$$P\Big(A > \theta E[A]\Big) \geq \frac{(1-\theta)^2 E[A]^2}{E[A^2]}$$

Chapter 4

Probabilistic Topic Modeling

The emergence of Big Data has made it more difficult than ever to organize and even identify the information and knowledge contained in large corpora of documents. Probabilistic topic modeling (PTM)[1] (Blei, 2012) is a form of *unsupervised learning* which is particularly suitable for unravelling information from massive sets of documents. Probabilistic topic models essentially consist of statistical methods that analyze words of documents and infer the probability with which each word of a given "vocabulary" is part of a *topic*. Intuitively, the words that have high probability in a topic, are very likely to appear together in a document that refers to this topic with high probability. Therefore, a topic, which is essentially a probability distribution of the words of a given vocabulary, reveals the underlying *hidden structure*. One of the most popular PTM algorithms (Blei, 2012) is *online Latent Dirichlet Allocation* (online LDA) (Hoffman, Bach, and Blei, 2010), which, as its name indicates, is a an online version of the well-known *Latent Dirichlet Allocation* (LDA) (Blei, Ng, and Jordan, 2003) algorithm. LDA is a generative probabilistic model for sets of discrete data, while online LDA can handle documents that arrive in streams, enabling the continuous evolution of the topics. Furthermore, LDA can be interpreted as a graphical model (Jordan, 1998; Jordan et al., 1999; Koller and Friedman, 2009) and a multinomial Principal Component Analysis (PCA) procedure (Hoffman, Bach, and Blei, 2010; Jolliffe, 2002; Abdi and Williams, 2010). Other notable PTMs are the mixture of unigrams model (Nigam et al., 2000) and the probabilistic latent semantic indexing model (Hofmann, 1999),

In this section, we provide the background that is essential for Chapter 6, where we present a novel method for agent-decision making in overlapping coalitional settings which is based on PTM, and specifically online LDA. Therefore, we begin by introducing LDA and then proceed in presenting online LDA.

[1]We may use the abbreviation PTM to refer to either Probabilistic Topic Modeling or Probabilistic Topic Model.

4.1 Latent Dirichlet Allocation

Since the field of Probabilistic Topic Modeling has emerged from applications in document analysis, the terminology that is usually used stems from that. However, the applications of PTM are much broader. In particular, the spectrum of domains in which Latent Dirichlet Allocation has been applied to ranges from bioinformatics (Liu et al., 2010) to music harmonic analysis (Hu and Saul, 2009) and software engineering (Maskeri, Sarkar, and Heafield, 2008).

Now, we define basic terms and notation:

- A *word* is the basic unit of discrete data. A vocabulary consists of words and is indexed by $\{1, 2, \ldots, V\}$, while it is fixed and has to be known to the LDA.
- A *document* is a series of L words, denoted by $\boldsymbol{w} = (w_1, w_2, \ldots, w_L)$, where the l^{th} word is denoted by w_l.
- A *corpus* is a collection of M documents, denoted by $D = \{\boldsymbol{w_1}, \boldsymbol{w_2}, \ldots, \boldsymbol{w_M}\}$.
- A *topic* is a distribution over a vocabulary.

LDA is a Bayesian probabilistic model, the intuition behind it being that a document is a mixture of topics. For each document $\boldsymbol{w}$ in D, LDA assumes the following generative process:

1. Pick in a random way a distribution over topics.
2. For each word in the document proceed in two steps:
   (a) Pick in a random way a topic from the distribution that was picked in step 1.
   (b) Pick in a random way a word from the topic that was picked in (a).

As implied by this generative process, documents share the same set of topics, but each exhibits topics in different portions.

While LDA observes only documents, i.e. sequences of words, its objective is to discover the *topic structure* which is *hidden*. It is thus assumed that the generative process includes latent variables. The topics are $\beta_{1:K}$, where K is their number; each topic β_k is a distribution over the vocabulary, where $k \in \{1, \ldots, K\}$; and β_{kw} is the probability of word w in topic k. For the d^{th} document the topic proportion of topic k is θ_{dk}, as θ_d is a distribution over the topics. The topic assignments for the d^{th} document are denoted by z_d, with z_{dl} being the topic assignment for the l^{th} word of the d^{th} document. Thus, β, θ and z are the latent variables of the model, while the only observed variable is w, where w_{dl} is the l^{th} word that is observed in the d^{th} document. Given the documents, the posterior of the topic structure is:

$$P(\beta_{1:K}, \theta_{1:D}, z_{1:D} \mid w_{1:D}) = \frac{P(\beta_{1:K}, \theta_{1:D}, z_{1:D}, w_{1:D})}{P(w_{1:D})}$$

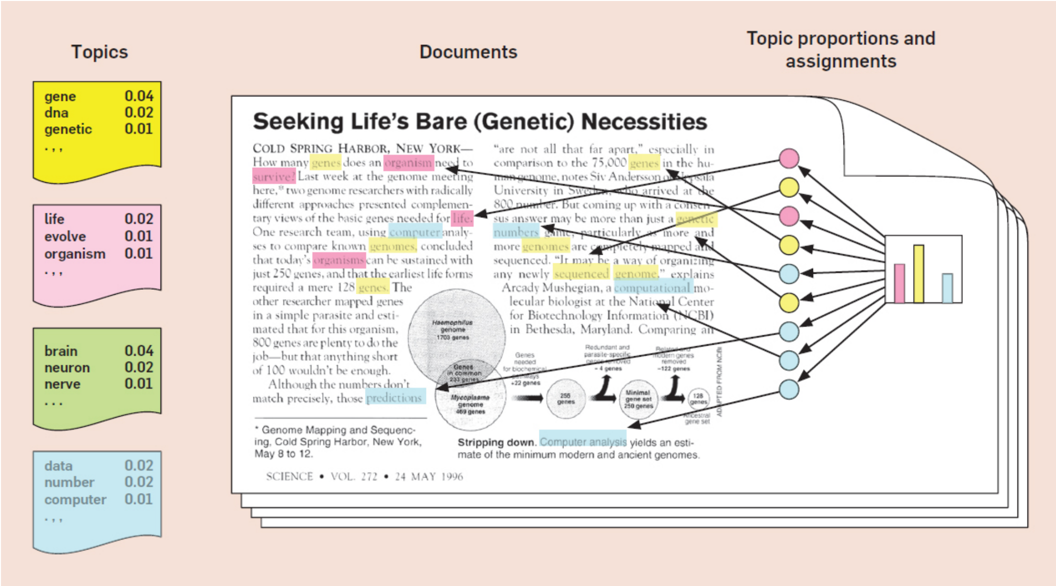


FIGURE 4.1: The intuitions behind Latent Dirichlet Allocation (Blei, 2012). Every document exhibits the same topics, but in different portion, and every topic exhibits the same words, but in different portion. Here, there are four topics (left), and the illustrated document is a distribution over these topics (right).

where the computation of $P(w_{1:D})$, the probability of seeing the given documents under any topic structure, is intractable (Blei, Ng, and Jordan, 2003; Blei, 2012). The difficulty of performing this computation stems from the coupling between θ and β. Furthermore, LDA introduces priors, so that $\beta_k \sim Dirichlet(\eta)$ and $\theta_d \sim Dirichlet(\alpha)$. The graphical model representation of LDA is depicted in Figure 4.2.

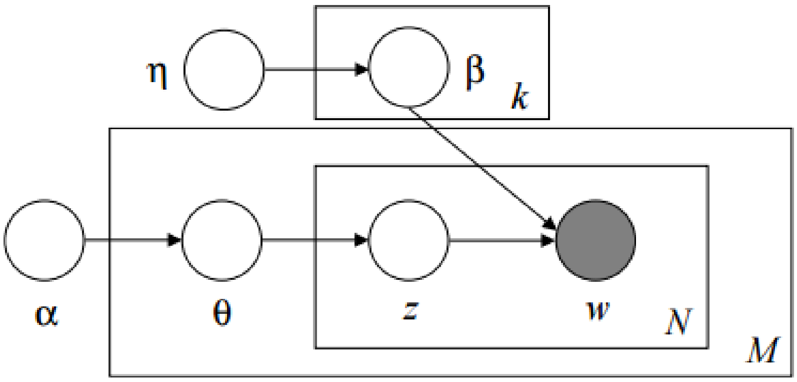


FIGURE 4.2: Graphical model representation of Latent Dirichlet Allocation (Blei, Ng, and Jordan, 2003).

Though the exact computation of the posterior, and thus the topic structure as a whole, cannot be efficiently computed, it can be approximated (Blei, Ng, and Jordan, 2003). The two most prominent alternatives for this are Markov Chain Monte Carlo (MCMC) sampling methods (Jordan, 1998) and variational inference (Jordan et al., 1999).

In variational inference for LDA, the true posterior is approximated by a simpler distribution q that depends on parameters (matrices) $\phi_{1:D}$, $\gamma_{1:D}$ and $\lambda_{1:K}$, defined as follows:

$$\phi_{dwk} \propto exp\{E_q[log\ \theta_{dk}] + E_q[log\ \beta_{kw}]\},$$
$$\gamma_{dk} = \alpha + \sum_w n_{dw}\phi_{dwk}, \quad \lambda_{kw} = \eta + \sum_d n_{dw}\phi_{dwk}$$

The variable n_{dw} is the number of times that word w has been observed in document d. Parameters $\gamma_{1:D}$ and $\lambda_{1:K}$ are associated with n_{dw}, while ϕ_{dwk} denotes the probability (under distribution q) that the topic assignment of word w in document d is k (Blei, Ng, and Jordan, 2003).

The variational inference algorithm (depicted in Algorithm 1) performs by minimizing the *Kullback-Leibler (KL) divergence* between the variational distribution and the true posterior. This is achieved via an *Expectation Maximization (EM)* procedure, where the algorithm iterates between assigning values to document-level variables and updating topic-level variables.

Algorithm 1: Variational Inference for LDA (Blei, Ng, and Jordan, 2003).

```
Initialize λ randomly
repeat
    E step :
    for d = 1 to D do
        Initialize γ_dk = 1 (or any other arbitrary constant)
        repeat
            Set φ_dwk ∝ exp{E_q[log θ_dk] + E_q[log β_kw]}
            Set γ_dk = α + Σ_w n_dw φ_dwk
        until (1/K) Σ_k |change in γ_dk| < ε
    M step :
    Set λ_kw = η + Σ_d n_dw φ_dwk
until relative KL divergence has not significantly decreased
```

4.2 Online Latent Dirichlet Allocation

In online LDA (Hoffman, Bach, and Blei, 2010), documents can arrive in batches (streams), and the value of $\lambda_{1:K}$ is updated through analyzing each batch of documents. The variable $\rho_t = (\tau_0 + t)^{-\kappa}$ controls the rate at which the documents of batch t impact the value of $\lambda_{1:K}$. Assigning a value to κ such that $\kappa \in (0.5, 1]$ guarantees the convergence of the algorithm. Furthermore, the algorithm requires an estimation, at least, of the total number of documents, D, in case this is not known in advance. The values of α and η can be assigned once and remain fixed. Essentially, the probability of word w in topic β_k, can be estimated as $\beta_{kw} = \lambda_{kw} / \sum \lambda_k$. Online LDA (using variational inference) is shown in Alg. 2 below.

Furthermore, a common means in document modeling for measuring the performance of the model is *perplexity*, which is computed over a held-out set of documents. Perplexity is estimated as the geometric mean of the inverse marginal probability of the held-out set of documents, and *the lower the perplexity the better the generalization performance* of the model (Blei, Ng, and Jordan, 2003). Formally, for a test set of documents D, where their number is M, the perplexity is:

$$perplexity(D_{\text{test}}) = \exp\left\{ -\frac{\sum_{d=1}^{M} log\, P(\boldsymbol{w}_d)}{\sum_{d=1}^{M} N_d} \right\}$$

where N_d is the number of words of document d belonging in the test set.

Algorithm 2: Online Variational Inference for LDA (Hoffman, Bach, and Blei, 2010).

Initialize λ randomly
for t = *1 to* ∞ **do**
    $\rho_t = (\tau_0 + t)^{-\kappa}$
    E step :
    Initialize γ_{tk} randomly
    repeat
        Set $\phi_{twk} \propto exp\{E_q[log\, \theta_{tk}] + E_q[log\, \beta_{kw}]\}$
        Set $\gamma_{tk} = \alpha + \sum_w n_{tw} \phi_{twk}$
    until $\frac{1}{K} \sum_k |\textit{change in } \gamma_{tk}| < \epsilon$
    M step :
    Compute $\tilde{\lambda}_{kw} = \eta + D n_{tw} \phi_{twk}$
    Set $\lambda = (1 - \rho_t)\lambda + \rho_t \tilde{\lambda}$

Chapter 5

Probability Bounds for Overlapping Coalition Formation

This chapter describes the first of the two main parts of this thesis. The main contribution is the development of three novel methods which benefit from obtaining probability bounds for assessing the ability of teams of agents to accomplish coalitional tasks. As argued before, in many realistic settings it is natural that individual agents contribute some amount of their (divisible) resource in order to complete the coalitional task. Therefore, we allow agents to join a number of *overlapping coalitions*, which are also referred to as *partial* coalitions (Chalkiadakis et al., 2010). Moreover, interconnected electronic societies (Papadimitriou, 2001) and social networks offer a natural environment for the completion of goals of individuals, most of the times by completing tasks through common actions. In this setting, the formation of a coalition is feasible only if its members are interconnected (Myerson, 1977).

Moreover, it is conceivable that an agent is certain only of the quantity of her own resource, while she has private beliefs about others' potential contribution. Thus, acting under incomplete information, she can only reason in a probabilistic manner about the success of a potential coalition (Chalkiadakis, Markakis, and Boutilier, 2007; Kraus, Shehory, and Taase, 2003). Against this background, we provide three novel methods that allow agents to establish probability bounds on the corresponding uncertainty. The bounds are derived given private agent beliefs about others' potential resource investment. These beliefs correspond to *Beta* and *Dirichlet* distributions which can be easily manipulated and updated in a principled manner. Our first method exploits an improvement of the *Paley-Zygmund inequality* (Paley and Zygmund, 1932), while the second and the third proceed by appropriately handling the *two-sided Chebyshev's inequality* (Mitzenmacher and Upfal, 2005) and the *Hoeffding's inequality* (Hoeffding, 1963), respectively. Agents (where the their set is denoted by $N = \{1, \ldots, n\}$) using any of them can demand an arbitrary *confidence level* for the resource contribution of a (partial) coalition.

In order for agents to complete tasks by forming overlapping coalitions in a decentralized manner, an appropriate protocol is required. Thus, we provide a generic protocol for iterated overlapping coalition formation. Under this protocol, in each iteration (round) a number of tasks arrives, and a proposer gets assigned to each task, who is responsible for forming a coalition so that the task gets completed; naturally, the underlying graph structure influences the plausibility of emergence of the coalitions. Though it suits the case of overlapping coalition formation, the protocol is not specific to it. To the best of our knowledge, this is the first paper to compute probability bounds for coalition formation under uncertainty.

We evaluate our methods by conducting experiments over both a 300 nodes Erdős-Renyi random graph and a real social network which is a 4039 nodes snapshot from Facebook (Leskovec and Krevl, 2014). Our results show that our methods consistently outperform, in terms of effectiveness in task completion, a baseline method that selects coalitions based only on expected resource quantity. Moreover, they are also time-efficient, and their behaviour is robust against increases in demanded confidence level. As such, we can conclude that they are indeed suitable for probabilistic reasoning in order to form coalitions in large-scale networks.

The focus of our work in this paper is not game-theoretical, nor do we attempt to conduct optimal coalition structure generation (Chalkiadakis, Elkind, and Wooldridge, 2011; Rahwan et al., 2015). Instead, we are interested in providing methods which, when employed by the agents, result in the effective formation of overlapping coalitions. We note that although we allow overlapping coalitions to form (in order to approach more realistic scenarios), our methods can be directly utilized by agents who operate in environments where no overlapping coalitions are plausible, also.

5.1 Modeling Uncertainty

The concise representation scheme of (Ueda et al., 2011), which is based on the idea of *agent types* (where agents $i, j \in N$ are recognizably equivalent if for any coalition C, $i, j \notin C : v(C \cup i) = v(C \cup j)$) cannot apply as a scheme for conciseness in threshold task games, as a task is characterized by a scalar, which is the sum of the resources required for its completion, and not by type vectors (as in (Ueda et al., 2011)). However, we achieve representational conciseness by restricting the quantities of the agent resources and the threshold values of the task types to integer values. Without these restrictions, the number of possible partial coalitions would be infinite. We thus assume that resources are integers, and an agent $i \in N$ possesses a resource amount $w_i \in \{1, \ldots, q_{max}\}$, where $q_{max} \in \mathbb{N}^+$.

5.1.1 Agent Beliefs

Each agent i knows with certainty only the quantity of her own resource, while she has a private belief X_j^i about the resource investment of each other agent j. Belief X_j^i is composed of random variables D_j^i and B_j^i, which are distributed as follows:

- $D^i_j \sim Multinomial\big(p^{ij} = (p^{ij}_1, \ldots, p^{ij}_{q_{max}})\big)$, where p^{ij}_r is the probability that j offers quantity r of her resource in a coalition.
- $p^{ij} \sim Dirichlet(\alpha^{ij}), \quad \alpha^{ij} = (\alpha^{ij}_1, \ldots, \alpha^{ij}_{q_{max}})$
- $B^i_j \sim Binomial(q^{ij})$, where q^{ij} is the probability that j accepts an offer by i for participation in a coalition.
- $q^{ij} \sim Beta(a^{ij}, b^{ij})$, where a^{ij} (b^{ij}) corresponds to the number of proposals from i that j has accepted (declined).

Note that α^{ij}, a^{ij} and b^{ij} above are hyperparameters corresponding to easily updated counters (as we explain later). We estimate p^{ij}_r as $\alpha^{ij}_r / \sum \alpha^{ij}$, and q^{ij} as $a^{ij}/(a^{ij} + b^{ij})$. Thus, $E[X^i_j] = q^{ij} \sum_{r=1}^{q_{max}} p^{ij}_r \cdot r$. Therefore, the expected, according to i's beliefs, quantity of the sum of resource contribution of a group of agents $X^i_C = \sum_{j \in C} X^i_j, C \subseteq N \setminus i$, can be computed in $O(n \cdot q_{max})$ time, through linearity of expectations. We assume that $X^i_1, \ldots, X^i_{i-1}, X^i_{i+1}, \ldots X^i_n$ are independent. Due to independence, the variance of X^i_C is the sum of the variances of each $X^i_j, j \in C$, and thus it can be computed in $O(n \cdot q_{max})$ time as well.

5.2 Obtaining and Exploiting Probability Bounds

Agents act under uncertainty regarding the amount of resources that the other agents may contribute to a coalition, for the completion of a common task, in the context of a threshold task game. In this section we describe how probability bounds over the resource contribution of a group of agents can be obtained, and how these can be exploited.

We henceforth refer to X^i_C as X, for fixed $i \in N$ and $C \subseteq N \setminus i$, when this causes no confusion. The threshold value T of a task $t = (T, u)$ must be exceeded, so that utility u is granted to the members of the successful coalition, and thus agent i is interested in computing $P(X \geq T)$. Since the cost of this computation can be extremely high,[1] the agents resort to computing bounds over this quantity. In the following subsections, we also provide reminders (as they were previously presented in section 3) of the definitions of the exploited probability inequalities.

[1]The distribution of the sum of independent integer-valued random variables can be computed via the *Convolution Theorem*, which exploits *Fast Fourier Transform* (Proakis and Manolakis, 1996), in $O(n^2 \cdot q_{max} \cdot \log(n \cdot q_{max}))$ time. However, this is pseudopolynomial, as it is polynomial in the numeric value q_{max}, and also quadratic in n (where we actually care about large n), and thus this computation time can be prohibitive. By contrast, we obtain (lower) probability bounds in $O(n \cdot q_{max})$ time, since the inequalities we utilize depend on the expected value and variance of X.

5.2.1 IPZY Method

Our first method builds on an **i**mprovement of the **P**aley-**Zy**gmund inequality (IPZY).

Theorem 6 (Paley-Zygmund inequality)**.** *Let a random variable $X \geq 0$ with finite variance and $0 \leq \theta \leq 1$. It holds:*

$$P(X > \theta E[X]) \geq \frac{(1-\theta)^2 E[X]^2}{E[X^2]} \tag{5.1}$$

Since $E[X^2] = Var[X] + E[X]^2$, it follows that:

$$P(X > \theta E[X]) \geq \frac{(1-\theta)^2 E[X]^2}{Var[X] + E[X]^2} \tag{5.2}$$

We now state and prove the following inequality that provides a better bound:[2]

Theorem 7. *Let a random variable $X \geq 0$ with finite variance, and $0 \leq \theta \leq 1$. It holds that:*

$$P(X > \theta E[X]) \geq \frac{(1-\theta)^2 E[X]^2}{Var[X] + (1-\theta)^2 E[X]^2} \tag{5.3}$$

Proof. The proof exploits the one-sided Chebyshev's inequality which states that for $\eta > 0$ the following two inequalities hold:

$$P(X \geq E[X] + \eta) \leq \frac{Var[X]}{Var[X] + \eta^2} \tag{5.4}$$

$$P(X \leq E[X] - \eta) \leq \frac{Var[X]}{Var[X] + \eta^2} \tag{5.5}$$

We proceed by exploiting Eq. (5.5), where the referred variables are defined as above, as follows:

$$P\big(X \leq \theta E[X]\big) = P\big(X \leq E[X] - (1-\theta)E[X]\big) \leq \frac{Var[X]}{Var[X] + (1-\theta)^2 E[X]^2}$$

[2]This improvement is reported at some places on the Web; to the best of our knowledge, we are the first to provide a formal proof.

It holds that:

$$
\begin{aligned}
& P\big(X > \theta E[X]\big) + P\big(X \leq \theta E[X]\big) = 1 \\
\Leftrightarrow\ & P\big(X > \theta E[X]\big) = 1 - P\big(X \leq \theta E[X]\big) \\
\Leftrightarrow\ & P\big(X > \theta E[X]\big) \geq 1 - \frac{Var[X]}{Var[X] + (1-\theta)^2 E[X]^2} \\
\Leftrightarrow\ & P\big(X > \theta E[X]\big) \geq \frac{Var[X] + (1-\theta)^2 E[X]^2 - Var[X]}{Var[X] + (1-\theta)^2 E[X]^2} \\
\Leftrightarrow\ & P\big(X > \theta E[X]\big) \geq \frac{(1-\theta)^2 E[X]^2}{Var[X] + (1-\theta)^2 E[X]^2}
\end{aligned}
$$

□

Clearly, Eq. (5.3) provides a better bound than Eq. (5.2), since for lower bounds greater values are preferred, and the denominator of Eq. (5.3) is smaller than the one of Eq. (5.2). Hence, lower probability bounds can be obtained via the employment of Eq. (5.3). Figure 5.1 presents typical values of the lower probability bounds, according to the beliefs of a random agent in our Erdős-Renyi experimental setting described later. Figure 5.1 also presents the corresponding values derived by Eq. (5.1) $\big($which provides the same results as Eq. (5.2)$\big)$, for comparison.

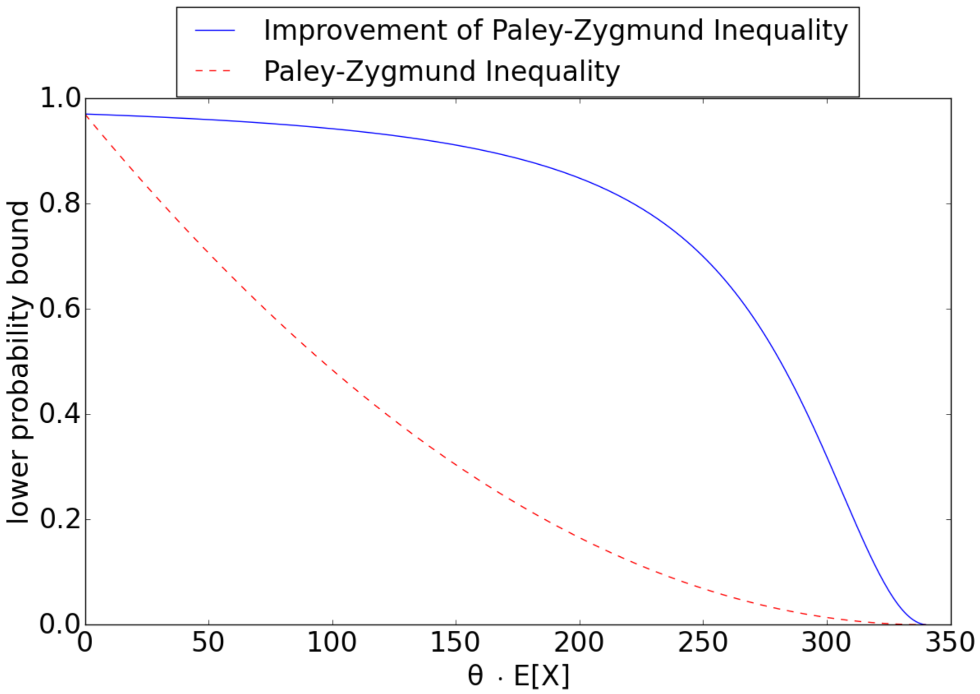


FIGURE 5.1: Lower bounds provided by Eq. (5.3) and Eq. (5.1) $\big($which provides the same results as Eq. (5.2)$\big)$, for a group of agents of size 45, with expected value 340.72 and variance 3529.71, based on the beliefs of a randomly selected agent.

Agent i wants to assess the ability of a group of agents C to complete task $t = (T, u)$. Thus, a lower bound for $P(X > T)$ can be obtained using Eq. (5.3), by setting:

$$\theta = T/E[X] \tag{5.6}$$

Notice that it must hold that $E[X] > T$, or else $\theta \geq 1$.

Furthermore, let i demand a certain *confidence level* c, $0 < c < 1$, so that the bound obtained for a group, using Eq. (5.3), with θ set as in Eq. (5.6), is greater than c. To select a group of agents, either to directly join or invite its members to join forces together, i would select the group that has *the smallest size* among those that offer her confidence level of at least c. In this way, i would have her confidence requirement satisfied; and by selecting an appropriate group with the smallest possible size, a greater portion of u could be distributed to each individual in the coalition.

5.2.2 CH2S Method

Our second method builds on a manipulation of the two-sided Chebyshev's inequality, and it is thus termed *CH2S*.

Theorem 8 (Two-sided Chebyshev's inequality). *Let a random variable X with finite variance, for any $k > 0$:*

$$P(|X - E[X]| \geq k) \leq \frac{Var[X]}{k^2} \tag{5.7}$$

Now, we can manipulate Eq. (5.7), in order to obtain a *lower probability bound*, in the following way:

$$P(|X - E[X]| < k) \geq 1 - \frac{Var[X]}{k^2} \tag{5.8}$$

The probability bound obtained by Eq. (5.8) can be exploited in a similar way as in the IPZY method, so we adopt the same notation as before. Here again, agent i wants to assess the ability of a group C to complete task $t = (T, u)$, and the lower bound of Eq. (5.8) can be obtained, by setting:

$$k = E[X] - T \tag{5.9}$$

Thus, i can obtain a lower bound, equal to $1 - Var[X]/k^2$, for the probability of the event that $X \in (T, E[X] + k)$, as illustrated in Figure 5.2. Agent i can select a group of agents, out of potentially many, as follows. She chooses the one that has the smallest size among those which provide her with a bound, computed by Eq. (5.8), and with k as in Eq. (5.9), that exceeds her required confidence level c, $0 < c < 1$.

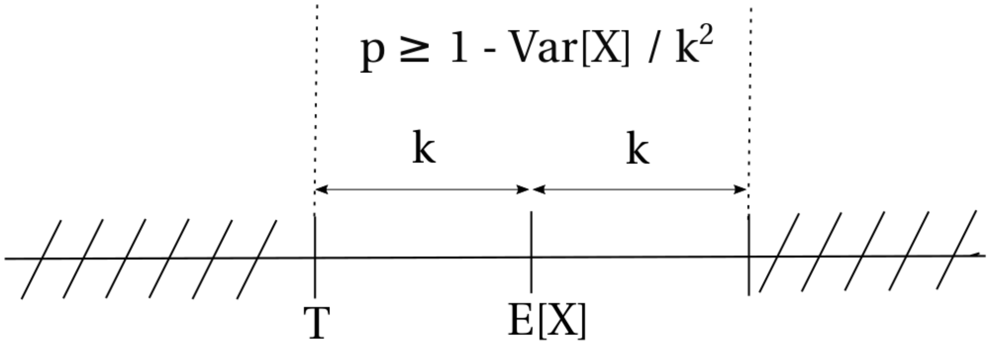


FIGURE 5.2: Illustration of Eq. (5.8)

5.2.3 HF Method

The third method (HF) is based on Hoeffding's inequality.

Theorem 9 (Hoeffding's inequality)**.** *Let $X_1, \ldots, X_n$ be independent random variables, where all X_i are bounded so that $X_i \in [l_i, u_i]$, and let $X = \sum_{i=1}^{n} X_i$. Then it holds that:*

$$P(|X - E[X]| \geq k) \leq 2exp\Big(-\frac{2k^2}{\sum_{i=1}^{n}(u_i - l_i)^2}\Big) \tag{5.10}$$

Hoeffding's inequality can be applied since random variables X_j^i are independent and X_j^i takes integer values in $[0, q_{max}]$. *HF* exploits Eq. (5.10) identically to the manner that $CH2S$ exploits Eq.(5.7), since both Eq.(5.7) and Eq.(5.10) provide upper bounds on both tails, and thus their difference is only on the value of the obtained bound.

5.2.4 Discussion

The power of our methods lies in the fact that they do not need to compute the exact distribution of X, since they obtain probability bounds. We have assumed variable independence, and thus an agent's beliefs can be represented by easy-to-maintain conjugate priors (Betas and Dirichlets) over these variables (Grimmett and Stirzaker, 2001). However, our methods do not depend in any way on the exact priors used.

Note also that $IPZY$ and $CH2S$ do apply if *dependence* among the variables holds, as well. In that case, computing the distribution of X would require

Algorithm 3: Coalition selection

Data: Agent's confidence level $c \in (0, 1)$
Input: A set of coalitions **C** $= \{C_1, C_2, \ldots\}$;
Task Threshold T
Output: Coalition $Best \in$ **C**

$min \leftarrow +\infty$, $Best \leftarrow \emptyset$, $bound \leftarrow 0$
foreach $C \in$ **C** **do**
    if *method* = $IPZY$ **then**
        $\theta \leftarrow T/E[X]$
        $bound \leftarrow \frac{(1-\theta)^2 E[X]^2}{Var[X]+(1-\theta)^2 E[X]^2}$
    else
        $k \leftarrow E[X] - T$
        if *method* = $CH2S$ **then**
            $bound \leftarrow 1 - \frac{Var[X]}{k^2}$
        else // $method = HF$
            $bound \leftarrow 1 - 2exp(-2k^2/(|C|q_{max}^2))$
    if $bound > c \wedge |C| < min$ **then**
        $Best \leftarrow C$
        $min \leftarrow |Best|$
return $Best$

the computation of the joint probability distribution of the dependent variables, leading to the following computation for $P(X_C^i \geq T)$ and $m = |C|$:

$$P(X_C^i \geq T) = \sum_{\kappa=T}^{m \cdot q_{max}} P(X_C^i = \kappa) = \sum_{\kappa=T}^{m \cdot q_{max}} P\big(\sum_{j=1}^{m} X_j^i = \kappa\big)$$

where the first equality holds since the events are mutually exclusive. The number of solutions to the equation $X_1^i + \ldots + X_m^i = \kappa$, where every X_j^i is a non-negative integer, is known (Komatsu, 2003) to be $\binom{\kappa+m-1}{m-1} = O((\kappa+m)^m)$. Therefore, the computation of $P(X \geq T)$ is exponential in the number of the agents when X_j^i, $j \in C$, are dependent. However, $IPZY$ and $CH2S$ would still work though requiring more time, since computing the variance of X would now take $O(n^2 \cdot q_{max})$, due to the computation of the covariance of each pair of variables.[3] Experimentation with dependent variables is future work.

One issue that arises regards the decision of an agent when her confidence level is *not* satisfied. This can occur in cases where either the cardinality of the set of coalitions to choose from is very small (i.e., the number of coalitions to pick from is very limited), or when the required confidence level c has a very high value. For instance, it is natural to expect that an agent with

[3] Furthermore, a different modeling of beliefs as Dirichlets over the Cartesian product of all variables, or perhaps as Dirichlet mixture models, would also be required.

$c = 1 - \epsilon$ (with ϵ very close to 0) will probably not manage to exceed it. In our experimental setting, we let the agents who did not manage to satisfy their confidence level to choose the coalition which provided the confidence level closest to c. Though this is natural, considering different alternatives is interesting future work also.

5.3 An Overlapping Coalition Formation Protocol

We now propose a generic protocol for iterated overlapping coalition formation. It models real-world situations where S tasks per iteration arrive, over a period of I iterations (or rounds), and the resources of the agents are *replenished* at the end of each round (as such, there is no need for long-term strategic planning on the part of the agents). In each round a proposer cannot be assigned with more than one task. Here we assume an underlying graph, however the protocol can be applied in environments where this is not the case. The agents of a coalition C must be connected by some path, however not all agents connected by a given path have to be members of C. The protocol is purposely kept simple, so that it does not interfere with the agents' deliberations over bounds. Despite it not being our most important contribution in this paper, we note that this is in fact the first generic protocol for decentralized overlapping coalition formation under uncertainty.

Indeed, while a protocol for task allocation is presented in (Weerdt, Zhang, and Klos, 2007) it disregards uncertainty and no overlapping coalitions can be formed. Similarly, protocols in (Kraus, Shehory, and Taase, 2003; Kraus, Shehory, and Taase, 2004) involve formation via setting-specific negotiations and auctions, and do not give rise to overlapping coalitions. To the best of our knowledge, protocols for decentralized *overlapping* coalition formation have only been developed in the field of telecommunications (Wang et al., 2014), where they are mostly domain-oriented and disregard uncertainty on the contribution of the agents.

The motivation behind our protocol's design lies on the way that research projects are assigned to institutions (e.g. universities, corporations with R&D departments), and the emerging need for cooperation among partners, so that the project gets successfully completed. Naturally, an institution (corresponding to an agent) can be simultaneously involved in a number of projects. Thus, it is natural to consider overlapping coalitions. Furthermore, not every institution can offer the same number of resources that are needed for the completion of a task, while others have (different) beliefs for one's contribution. Assuming an underlying social network is natural for representing the plausibility of cooperation among the institutions. However, our protocol is not domain-specific, and can be straightforwardly employed for other suitable applications.

We now present the details of the protocol. At each round, tasks are exogenously created and a proposer is exogenously associated with each task.

Since the choice of the proposers has to be non-deterministic, their assignment to tasks takes place uniformly at random. Right after the assignment of proposers to tasks, the proposers are concurrently asked to form proposals to be submitted to other agents. Therefore, each proposer has to select a group of agents, offering a portion of u (the utility granted by the completion of task $t = (T, u)$) to each of its members, and asking in return for a resource quantity so that the threshold T is exceeded and thus have the task completed.

5.3.1 Group Selection

The number of possible groups of agents is $O(2^n)$, and thus taking all of them under consideration would be inefficient for a proposer. Instead, we let a proposer i *sample* K groups of agents, where each agent $j \in N \setminus i$ is included in a group with probability $2^{-D(i,j)}$, where $D(i, j)$ is the geodesic distance (shortest path) between nodes i and j. Such a distance between every pair of nodes can be computed in an offline step. Hence, the closer i is to j, with respect to their distance in the (undirected) graph, the more likely it is that j will be included in a group sampled by i, while if i and j are not connected, that is $D(i, j) = \infty$, then they cannot cooperate. In this way, the position of an agent in the social network affects the likelihood of making a proposal to another agent, for cooperation towards completing a common task. Thereafter, the proposer uses one of our methods in order to select one of the K groups, and then submits individual $\langle q, \pi \rangle$ proposals to its members. The *requested quantity* $q \in \mathbb{N}^+$ is the rounded average of samples taken from α^{ij} multiplied by q^{ij} (the belief that j accepts a proposal), and *agent payoff* $\pi \in \mathbb{R}^+$ is proportional to q, and is distributed to j upon the completion of the task.

Only the proposer and the proposed-to agent have knowledge of the proposal submitted. Furthermore, all proposals to an agent are revealed to her simultaneously.[4]

5.3.2 Response to Proposals

Each agent, who has received at least one proposal, responds by either accepting or rejecting each of the proposals. Therefore, agent j has to select which of the, at most S, proposals of the form $\langle q, \pi \rangle$ to accept, where $q \in \mathbb{N}^+$ and $\pi \in \mathbb{R}^+$, in order to *maximize* her gained utility (at the completion of the tasks), given her resource quantity $w_j \in \mathbb{N}^+$. Thus, the optimal response of an agent is the one that maximizes $\sum_\tau x_\tau \cdot \pi_\tau$, subject to $\sum_\tau x_\tau \cdot q_\tau \leq w_j$, where x is a binary vector of size equal to the number of the proposals that agent j has received.

[4] Guaranteeing that the protocol in fact succeeds in revealing all proposals to an agent simultaneously, and ensuring that revelations are made to the members involved only, is a problem of broad interest in distributed and multi-agent systems, and can be approached by methods that are orthogonal to our work here. In particular, recent suggestions include the use of crypto-systems (Franco, 2014).

Proposition 1. *The optimal response of an agent is **NP**-Hard.*

Proof. This follows from a straightforward reduction from KNAPSACK (Garey and Johnson, 1979). Given ν items with size $s_1, s_2, \dots, s_\nu$, value $v_1, v_2, \dots, v_\nu$, and capacity B, where $s_\kappa \in \mathbb{N}^+, v_\kappa \in \mathbb{R}^+$, for $\kappa \in \{1, \dots, \nu\}$ and $B \in \mathbb{N}^+$, an item $\langle s_\kappa, v_\kappa \rangle$ corresponds to a proposal $\langle q, \pi \rangle$ and the capacity B to the resource quantity w_j of agent j. □

Notice that a responder has no knowledge of the other members that a proposer has proposed to, and hence she cannot infer the probability of the completion of the task. Moreover, since the response of an agent does not depend on the responses of the rest of the agents, the assumption on the independence of the variables related to agent beliefs (regarding the resource contribution of others), is reasonable.

5.3.3 Beliefs update

If the response of j to proposal $\langle q, \pi \rangle$, submitted by proposer i, is positive then i increases both a^{ij} (Beta update) and α_q^{ij} (Dirichlet update) by 1. If the response is negative then b^{ij} (only Beta update) is increased by 1. Furthermore, if coalition C succeeds in completing a task, then every member of C gets to learn the contribution of every other member, and has her Dirichlet distribution updated in an identical way to that of the proposer.[5]

Notice this allows the modeling of agent behaviour and preferences, e.g., an agent could observe that another agent accepts proposals from the rest of the agents but not from her. We aim to study such phenomena in future work.

Finally, after receiving the responses, if the task of a proposer has not been accomplished then she covers its remaining needs (if she is able to).

5.4 Experiments

In this section we provide results on the effectiveness of our methods. We conducted experiments on both an Erdős-Renyi random graph (Bollobás, 2001) of 300 nodes-agents and a real social network—a snapshot of a part of Facebook with 4039 agents (Leskovec and Krevl, 2014). Our methods are tested for different values of confidence level c demanded by the agents. We compare our methods to a baseline method, in which the group that an agent chooses to make proposals to is the *smallest* among those whose *expected value* $E[X]$ exceeds the task threshold value T.[6] Thus, we refer to this method

[5] Notice that in that case, the Beta distribution of an agent is not updated (because Beta models responses of others to our own proposals). We could have modeled the fact than an agent accepts the proposals of others but not ours, via the use of additional Beta.

[6] The smaller the group, the greater the portion of utility u its members receive upon task completion.

Algorithm 4: Overlapping Coalition Formation Protocol

$I \leftarrow$ number of rounds
$S \leftarrow$ number of tasks per round
for $i = 1$ *to* I **do**
  $tasks \leftarrow createTasks(S)$
  $proposers \leftarrow createProposers(tasks)$
  $proposers.formProposals()$ concurrently
  inform proposed-to members
  inform proposers about responses
  foreach *uncompleted* $t \in tasks$ **do**
    Let proposer of t complete it
  foreach *completed* $t \in tasks$ **do**
    inform members about the participation of the rest and distribute utility
  replenish the resources of the agents

Algorithm 5: formProposals

Sample K groups of agents
$Best \leftarrow$ group selected using *IPZY CH2S*, or *HF*
foreach $j \in Best$ **do**
  form a proposal $\langle q, \pi \rangle$ for j

as *EV*. Selecting EV as the baseline is natural: EV essentially describes the simplest course of action that an agent acting under uncertainty could take. Moreover, there is no pool of alternative decision-making methods we can use as a baseline in such a setting.[7]

Game parameters. In the experiments on both graphs, q_{max} was set to 30 and the resource weight w_i of each agent was a (rounded to integer) sample from $\mathcal{N}(15, 5^2)$. The hyperparameters of each agent's Beta were initialized to $a^{ij} = 1$ and $b^{ij} = 1$; and the α_r^{ij} of the Dirichlets to $\alpha_r^{ij} = w_j/(D(i,j) \cdot (|r - w_j| + 1))$. In this way α^{ij} is bell-shaped, with its mode being at w_j, the actual (maximum offerable) resource quantity of j. Furthermore, the smaller $D(i,j)$ is, the more sharply peaked the prior is, so the belief updates on α^{ij} have greater impact. The value of K, the number of groups that a proposer i samples, was set to 30. The number of samples taken from α^{ij}, for defining the requested quantity q of i's proposal $\langle q, \pi \rangle$ to j, was set to 20. The number of rounds I and the number of tasks per round S were set to 200 and 16, respectively,[8] for each

[7] The computation of a centralized solution is intractable, and thus it cannot be used as a means of comparison. Even the positive results in Zick, Chalkiadakis, and Elkind, 2012 hold for very restricted cases.

[8] In our experiments we used the standard pseudopolynomial dynamic programming algorithm for KNAPSACK (Kellerer, Pferschy, and Pisinger, 2004), which, due to the limited number of proposals, performed in a fraction of time ($< 1\ ms$), as expected (despite Proposition 1); so, agent responses were in fact optimal.

run on both graphs. Before making proposals, a proposer invests 75% of her resource to a task.

All results are average values across 30 runs for each experimental setting. The same set of proposers and tasks were generated at the x^{th} round of the y^{th} run in each setting ($x \in \{1, \ldots, I = 200\}$ and $y \in \{1, \ldots, 30\}$). We report on the total number of completed tasks by the agents, where for each setting the used method and the value of c were the same for all agents. We tested values of $c \in \{0.5, 0.6, 0.7, 0.8, 0.9\}$. (In the case of *EV* there is no c.) We also present the (average) *size of the group* that the proposers, using each method, choose as the *best* one to make proposals to. Furthermore, since there is no guarantee that the agents will have their demand for confidence level c fulfilled, we report the *confidence fails*, the number of times that a proposer failed to achieve that c. In those cases, we let the proposer make proposals to the group which provided the confidence level closest to c. Henceforth, we denote the *best group size* by *bgs*, and the *confidence fails* by *cf*. The total number of tasks in a run was $I \cdot S = 200 \cdot 16 = 3200$ (thus *cf* has a max value of 3200). The implementation was in Python 3 and experiments were run on a PC with an i3 3.3GHz processor and 4GB of RAM.

Algorithm 6: Dynamic Programming Algorithm for KNAPSACK

Input: n (number of items), W (weight of knapsack), c (value of items), w (weight of items)

Output: opt (the items that are included in the optimal solutions can be found via employing an additional matrix)

for $i = 0$ *to* W **do**
    $g[i, 0] \leftarrow 0$
for $j = 0$ *to* n **do**
    $g[0, j] \leftarrow 0$
for $i = 1$ *to* W **do**
    for $j = 1$ *to* n **do**
        if $w[j] > i$ **then**
            $g[i, j] \leftarrow g[i, j-1]$
        else
            $g[i, j] \leftarrow max\{g[i, j-1], c[j] + g[i - c[j], j-1]\}$
opt $\leftarrow g[W, n]$
return *opt*

5.4.1 Erdős-Renyi Graph Model

In this random graph model, edge $\{i, j\}$ is added on the (undirected) graph with probability p. In our setting, we set $n = 300$ and $p = 0.03$ (average node distance = 2.83). The threshold values T of the generated tasks were sampled from $\mathcal{N}(200, 10^2)$, and the utility u of each task was $10 \cdot T$.

As shown in Table 5.1, our methods consistently outperform *EV*. Even for *IPZY* with $c = 0.5$, the number of completed tasks is much larger compared to that of *EV*, with a very small increase in *bgs*. The standard deviation of the number of completed tasks is small for every method. Furthermore, when $c \leq 0.8$, proposers using $IPZY$ or $CH2S$ have their required confidence satisfied almost always. For $c = 0.9$ however, the number of *cf* increases steeply: for instance, for *CH2S* with $c = 0.8$, only $1.33\% = 42.7/3200$ of the proposals did not meet that requirement, while for *CH2S* with $c = 0.9$ that percentage increased to $33.33\% = 1066.8/3200$. Now, lack of confidence, caused by uncertainty, is a major source of downturns in economic environments (Caldara et al., 2016). Ideally, we would like to have a metric defining the quality of a method, depending on the number completed tasks and confidence fails, but it is not obvious which one that should be. Thus, HF can be deemed unsuitable for this setting, since the number of cf for every value of c is at least 1022.4, and so are $IPZY$ and $CH2S$ for $c = 0.9$ —notwithstanding that for $c = 0.9$ more tasks were completed, since their completion came with an abrupt increase in *cf*. However, for $c = 0.8$ for both $IPZY$ and $CH2S$ the highest confidence level that was experimentally shown to be satisfied almost always. Then, for $IPZY$ and $CH2S$, results presented in Table 5.3 suggest that the more demanding the agents are (with respect to their confidence level), the better the ratio of the number of completed tasks to best group size is. Overall, Tables 1 and 3 suggest that $CH2S$ perform slightly better than $IPZY$.

Finally, we report that the average time for an entire round to be completed, including coalitional evaluations for all 16 tasks and belief updates for all agents—for some multiple times, since they can participate in multiple tasks—was 2.7 sec for *IPZY*, 2.9 sec for *CH2S* and 1.8 sec for *HF*.

5.4.2 Facebook Snapshot Graph

The graph of the snapshot of Facebook we experimented on consists of 4039 connected nodes (agents). The threshold values T of the generated tasks were sampled from $\mathcal{N}(2250, 40^2)$, and the utility u of each task was $10 \cdot T$.

As observed in Table 5.2, task completion was more challenging in this sparser network (with an average node distance of 3.69). However, the ordering of the methods, for $IPZY$ and $CH2S$, based on the ratio of the number of completed tasks to *bgs*, was virtually the same with that of the Erdős-Renyi graph, as observed in Table 5.3. Moreover, we observe that proposers using HF completed more tasks compared to the other two methods, for every value of c, with the value of cf not being tremendously larger. The values in Table 5.3 corresponding to HF were greater compared to those of $IPZY$ and $CH2S$ for any value of c, and monotonically increase with c, even exceeding 3 for $c = 0.9$. Hence, our methods outperform the baseline in this

TABLE 5.1: Results for the Erdős-Renyi graph with 300 agents.

	#completed tasks (%) {std}	bgs	cf
EV	637.46 (19.92) {22.78}	35.16	-
IPZY, c = 0.5	876.70 (27.39) {18.50}	36.27	0.1
IPZY, c = 0.6	990.76 (30.96) {24.97}	36.87	0.2
IPZY, c = 0.7	1202.16 (37.56) {24.83}	38.13	1.6
IPZY, c = 0.8	1615.30 (50.47) {22.64}	41.30	14.9
IPZY, c = 0.9	2352.51 (73.52) {18.43}	52.95	692.9
CH2S, c = 0.5	1115.33 (34.85) {17.29}	37.60	1.1
CH2S, c = 0.6	1235.26 (38.60) {22.46}	38.40	2.6
CH2S, c = 0.7	1451.73 (45.36) {19.82}	39.94	6.9
CH2S, c = 0.8	1843.96 (57.62) {15.46}	43.52	42.7
CH2S, c = 0.9	2396.81 (74.91) {16.23}	54.85	1066.8
HF, c = 0.5	2316.83 (72.40) {21.63}	52.86	1022.4
HF, c = 0.6	2356.80 (73.65) {18.42}	54.95	1616.1
HF, c = 0.7	2368.80 (74.02) {11.66}	56.48	2213.9
HF, c = 0.8	2370.86 (74.09) {17.45}	57.46	2727.7
HF, c = 0.9	2372.03 (74.13) {18.29}	57.88	3095.6

TABLE 5.2: Results for the Facebook graph with 4039 agents.

	#completed tasks (%) {std}	bgs	cf
EV	462.26 (14.45) {31.16}	396.12	-
IPZY, c = 0.5	562.76 (17.58) {29.62}	398.68	241.1
IPZY, c = 0.6	608.31 (19.01) {28.16}	399.97	295.7
IPZY, c = 0.7	676.30 (21.13) {25.99}	401.68	397.5
IPZY, c = 0.8	803.50 (25.11) {26.95}	404.45	525.4
IPZY, c = 0.9	1080.33 (33.76) {24.07}	412.39	742.5
CH2S, c = 0.5	648.60 (20.26) {26.29}	400.48	356.8
CH2S, c = 0.6	684.61 (21.40) {28.92}	401.27	411.8
CH2S, c = 0.7	747.10 (23.35) {27.92}	402.61	482.8
CH2S, c = 0.8	860.93 (26.90) {25.56}	405.19	565.4
CH2S, c = 0.9	1108.21 (34.63) {24.57}	412.69	799.1
HF, c = 0.5	1147.66 (35.86) {11.71}	408.35	772.4
HF, c = 0.6	1171.20 (36.60) {15.72}	410.17	852.4
HF, c = 0.7	1201.93 (37.56) {16.42}	412.46	946.1
HF, c = 0.8	1242.40 (38.82) {12.31}	415.42	1071.1
HF, c = 0.9	1285.46 (40.17) {11.92}	420.02	1272.0

TABLE 5.3: # of completed tasks to best group size ratio.

	Erdős-Renyi graph	Facebook graph
EV	18.13	1.17
IPZY, c = 0.5	24.17	1.41
IPZY, c = 0.6	26.87	1.52
IPZY, c = 0.7	31.52	1.68
IPZY, c = 0.8	39.11	1.99
IPZY, c = 0.9	44.42	2.62
CH2S, c = 0.5	29.66	1.62
CH2S, c = 0.6	32.16	1.71
CH2S, c = 0.7	36.34	1.84
CH2S, c = 0.8	42.37	2.12
CH2S, c = 0.9	43.69	2.69
HF, c = 0.5	43.82	2.81
HF, c = 0.6	42.89	2.85
HF, c = 0.7	41.94	2.91
HF, c = 0.8	41.26	2.99
HF, c = 0.9	40.98	3.06

environment as well; and HF appears to be a winner.[9]

The difficulty in achieving the completion of tasks can be attested by the fact that even when proposers used the EV method, about $6.5\% = 207.2/3200$ of them could not find a (sampled) group whose expected value exceeded the task threshold value. In that case they selected the one with the highest expected value. Thus, it is not surprising that for all our three methods, and for every value of confidence level c, proposers frequently could not fulfill their requirement for exceeding c. The fact that *cf* increases smoothly with c, as seen in Table 5.2, combined with the significant increase of the ratio of completed tasks to best group size for $c = 0.9$, observed in Table 5.3, lets us conclude that the best value of c in this setting would be 0.9. The average time for the completion of an entire round was 26.5 sec for *IPZY*, 27.7 sec for *CH2S*, and 21.3 sec for *HF*.

5.5 Conclusions

We presented, for the first time in the literature, three methods that derive probability bounds for effective overlapping coalition formation, where the agents have incomplete information of the value of the resources that the other agents can invest. We also proved an improvement of the Paley-Zyg-

[9]The values of the ratios in Table 5.3 for the Facebook graph are much smaller than those for the Erdős-Renyi graph, since the number of the agents of a coalition is much greater compared to that in the Erdős-Renyi graph. In both graphs, however, proposers sampled groups in the way mentioned in the previous section.

mund inequality, and presented a protocol for overlapping coalition formation. Each of the proposed methods exploits a different probability inequality (improvement of the Paley-Zygmund's inequality, two-sided Chebyshev's inequality, and Hoeffding's inequality). This setting extends straightforwardly to environments with multiple (rather than one) resource types. All three allow agents to demand confidence levels; and significantly outperformed the baseline (which picked coalitions based solely on their expected resource quantity) in terms of the number of tasks completed, and the ratio of this quantity to the size of the group the proposers selected.

Chapter 6

Overlapping Coalition Formation via Probabilistic Topic Modeling

This chapter presents the second work that was conducted in the context of this thesis. The main contribution is the development of a completely novel approach in multi-agent learning for (overlapping) coalition formation, which is based on Probabilistic Topic Modeling, and, especially, online Latent Dirichlet Allocation. Furthermore, our contributions include the development of a concise representation scheme for cooperative games with overlapping coalitions which extends the much celebrated Marginal Contribution networks (MC-nets) scheme.

In many realistic environments (such as the one presented in the context of the previous chapter) agents are able to invest only a portion of their divisible resource in cooperating with others, and thus form overlapping coalitions. Therefore, an individual can participate in a number of coalitions simultaneously (Chalkiadakis et al., 2010). Additionally, as real-world environments exhibit a high level of uncertainty, it is more natural than not to assume that agents do not have complete knowledge of the utility that can be yielded by every possible team of agents (Suijs et al., 1999; Chalkiadakis and Boutilier, 2004; Kraus, Shehory, and Taase, 2003; Ieong and Shoham, 2008). Moreover, coalitional value is often determined by an underlying structure defined given *relations* among the members of the coalition (Ieong and Shoham, 2005). These relations reflect the synergies among the coalition members. It is natural to posit that agents do not know the exact synergies at work in their coalitions. Against this background, in our system the coalitional value depends on the amount of resources the agents invest, and, crucially, the explicit relations among coalition members. As such, we build on the idea of *Marginal Contribution nets* (MC-nets) (Ieong and Shoham, 2005) and introduce *Relational Rules* (RR), a representation scheme for cooperative games with *overlapping* coalitions. The RR scheme allows for the concise representation of the synergies-dependent coalition value.

Now, an agent can make an observation of the utility that can be earned by the resource offerings of the members of a coalition, but it is a much more complex task to determine her relations with subsets of agents of that coalition. Probabilistic topic modeling (PTM) (Blei, 2012) (presented in Chapter

4) consists a common approach in performing document analysis, and, in particular, in extracting the hidden thematic structure, in the form of topics (distributions over the words of a vocabulary), of the analyzed documents. In our approach, we exploit *online Latent Dirichlet Allocation* (online LDA) (Hoffman, Bach, and Blei, 2010), which can handle documents that arrive in streams, enabling the continuous evolution of the topics.

Our method for decentralized overlapping coalition formation employs online LDA to allow agents to learn how well they can cooperate with others. In our setting, agents *repeatedly* form overlapping coalitions, as the game takes place over a number of *iterations*. Therefore, we utilize a simple, yet appropriate, protocol, under which in each iteration an agent is (randomly) selected in order to propose (potentially) overlapping coalitions. Agents that use our method take decisions on which coalitions to offer some amount of their resource to, and which portion of that resource to each coalition, by exploiting the topics of the model that they have learned via employing online LDA: By interpreting formed coalitions as documents, represented given an appropriate vocabulary, agents are able to use online LDA to update beliefs regarding the hidden collaboration structure, and thus implicitly learn the most rewarding synergies with others (synergies described by RRs). Furthermore, agents are able to gain knowledge regarding the coalitions that are costly, and should thus be avoided. As a result, agents can, over time, pick partners with which to cooperate effectively.

To the best of our knowledge, this is the first time that Probabilistic Topic Modeling is employed for multi-agent learning (Tuyls and Weiss, 2012). In order to evaluate the performance of our method, we have developed a Q-learning (Watkins and Dayan, 1992) style algorithm, which we use as a baseline. Our algorithm vastly outperforms the baseline, implying a high degree of accuracy in the beliefs of the agents, and a high quality of agent decisions.

6.1 Relational Rules

In the real world agents often do not have complete knowledge of how well they can cooperate with others. An agent cannot be sure of the utility that will be earned by the formation of a coalition, since she has incomplete information regarding the efficiency with which its members can work together. Thus, agents have to form coalitions under what we term *structural uncertainty*. This notion describes the uncertainty agents face regarding the value of *synergies* among them.

Such synergies are, in a non-overlapping setting, concisely described by *Marginal Contribution networks* (MC-nets). In MC-nets, coalitional games are represented by a set of rules of the form :

$$Pattern \rightarrow value$$

where $Pattern$ is a conjunction of literals (representing the participation or absence of agents), and applies to coalition C if C satisfies $Pattern$, with $value \in \mathbb{R}$ being added to the coalitional value of C.

We now extend MC-nets to overlapping environments. We introduce *Relational Rules* (RR), with the following form:

$$\{i, j, \ldots, k\} \rightarrow \frac{\pi_{i,C} + \pi_{j,C} + \ldots + \pi_{k,C}}{|\{i, j, \ldots, k\}|} \cdot value$$

where $value \in \mathbb{R}$; $C \subseteq N$ (with $N = \{1, \ldots, n\}$ being the set of agents) is a coalition such that $\{i, j, \ldots, k\} \subseteq C$; $\pi_{i,C}$ is the portion of her resource that i has invested in coalition C: i.e., $\pi_{i,C} = r_{i,C}/r_i$, where r_i is the total resource quantity (continuous or discrete) that i holds and $r_{i,C}$ is the amount she has invested in C. Therefore, $\pi_{i,C} > 0$, since $i \in C$ ($r_{i,C} = 0$ essentially means that $i \notin C$), and $\pi_{i,C} \leq 1$, since the maximum resource offering of i to C is r_i.

A rule applies to coalition C if and only if $\{i, j, \ldots, k\} \subseteq C$, and in that case utility $\frac{\pi_{i,C} + \pi_{j,C} + \ldots + \pi_{k,C}}{|\{i,j,\ldots,k\}|} \cdot value$ is added to the coalitional value of C. Note that it is *not* required that an agent's r_i has to be communicated to C's other members, since a rule is applied by the environment. For non-overlapping games, RRs reduce to MC-nets rules without negative literals, as it then holds that $\frac{\pi_{i,C} + \pi_{j,C} + \ldots + \pi_{k,C}}{|\{i,j,\ldots,k\}|} = 1$.

Example 6.1.1. *Assume that* $N = \{1, 2, 3, 4\}$, $r_1 = 10, r_2 = 8, r_3 = 8, r_4 = 6$ *and the Relational Rules of the game are:*

$$\{1, 2, 3\} \rightarrow \frac{\pi_{1,C} + \pi_{2,C} + \pi_{3,C}}{3} \cdot 100 \tag{6.1}$$

$$\{2, 3\} \rightarrow \frac{\pi_{2,C} + \pi_{3,C}}{2} \cdot 80 \tag{6.2}$$

$$\{3, 4\} \rightarrow \frac{\pi_{3,C} + \pi_{4,C}}{2} \cdot (-100) \tag{6.3}$$

Furthermore, let coalitions $C_1 = \{1, 2, 3\}$ *and* $C_2 = \{2, 3, 4\}$ *form, with* $r_{1,C_1} = 10$, $r_{2,C_1} = 6$, $r_{3,C_1} = 4$, *and* $r_{2,C_2} = 2$, $r_{3,C_2} = 4$, $r_{4,C_2} = 6$. *Therefore,* $\pi_{1,C_1} = 10/10 = 1$, $\pi_{2,C_1} = 6/8 = 0.75$, $\pi_{3,C_1} = 4/8 = 0.5$, *and* $\pi_{2,C_2} = 2/8 = 0.25$, $\pi_{3,C_2} = 4/8 = 0.5$, $\pi_{4,C_2} = 6/6 = 1$.

The value u_{C_1} *of coalition* C_1 *will be determined by rules* (6.1) *and* (6.2), *since rule* (6.3) *does not apply as* $\{3, 4\} \nsubseteq C_1$. *Applying rule* (6.1) *to* C_1 *will result in value* $\frac{\pi_{1,C_1} + \pi_{2,C_1} + \pi_{3,C_1}}{3} \cdot 100 = (1 + 0.75 + 0.5)/3 \cdot 100 = 75$ *and applying rule* (6.2) *to* C_1 *will result in value* $\frac{\pi_{2,C_1} + \pi_{3,C_1}}{2} \cdot 80 = (0.75 + 0.5)/2 \cdot 80 = 50$. *Thus,* $u_{C_1} = 75 + 50 = 125$. *Following that reasoning, the value of coalition* C_2 *is determined by applying rules* (6.2) *and* (6.3), *which result to utilities* 30 *and* -75 *respectively, and thus* $u_{C_2} = 30 - 75 = -45$.

In our setting, the value of a coalition is determined through RRs, but agents *do not know the RRs in effect*, and hence cannot determine the value of a coalition with certainty. Thus, agents do not know how well they can do with others, and cannot determine their relations just by an observation of a coalitional value. However, in following section we show how PTMs can be exploited so that agents implicitly learn the underlying RR-described collaboration structure.

6.2 Learning by Interpreting Coalitions as Documents

In cooperative games , overlapping or not, it is natural for agents to take decisions regarding the formation of coalitions, and to update their beliefs based on subsequent observations. Naturally, an agent receives information only about the coalitions that have been formed and she is a member of. The decision-making process of the agents is addressed by our method, which is presented in the next section. In this section, we present how agents can employ online LDA in order to effectively learn the underlying collaboration structure. We let each agent maintain and train her own online LDA model. Thus, there are n such models in the system.

For each partial coalition $C \subseteq N$, $i \in C$, i observes the earned utility u_C. The contribution $r_{i,C}$ and the utility $u_{i,C}$, earned by i from participating in C, are known to each other agent $j \in C \setminus i$. However, in order to supply that information to her online LDA model, an agent must maintain a vocabulary. We define the vocabulary of an agent to include $n-1$ words, one for each other agent, indicating their contribution, plus two words for the utility, one representing gain and the other representing loss, since the value earned from a partial coalition can be either positive or negative. Therefore, the vocabulary of an agent consists of $n+1$ words.[1] Assuming a game that proceeds in rounds, in round t agent i interprets the coalitional configuration regarding $C, i \in C$, as a document by "writing" in the document the word that indicates the contribution of agent $j \in C \setminus i$ $\;r_{j,C}$ times—where $r_{j,C} \in \mathbb{N}^+$ is the contribution of j to C. The restriction of the resource contributions of agents to positive ($r_{j,C} = 0 \Rightarrow j \notin C$) natural numbers is thus necessary when LDA is used, since a word can only appear in a document a discrete number of times. Therefore, the number of documents that an agent passes in an iteration (round) to her online LDA is equal to the number of coalitions that she is member of. (Naturally, if an agent takes part in a coalition with the same configuration in a future iteration, then the same document will be formed and passed to her model.) Therefore, agent i "writes" in the document, that corresponds to C, either the word that indicates gain or the one that indicates loss as many times as the absolute value of the utility earned by the

[1] An obvious extension is to let an agent include a word for her own contribution in the vocabulary. Naturally, this would affect the agent's decision-making process.

coalition is.[2] Since words are discrete data, u_C cannot be real-valued; so, we let the actual value earned by C be $\lfloor u_C \rfloor$, instead of the u_C computed by the application of the RRs related to C.

Example 6.2.1. *Let agent* 1*'s vocabulary include the words "ag2", "ag3", "gain" and "loss", corresponding respectively to agents'* 2 *and* 3 *contribution and the positive and negative utility. Therefore, agent* 1 *forms, for coalition* $C = \{1, 2, 3\}$ *where* $r_{2,C} = 1$, $r_{3,C} = 2$ *and* $u_C = -4$, *the document:*

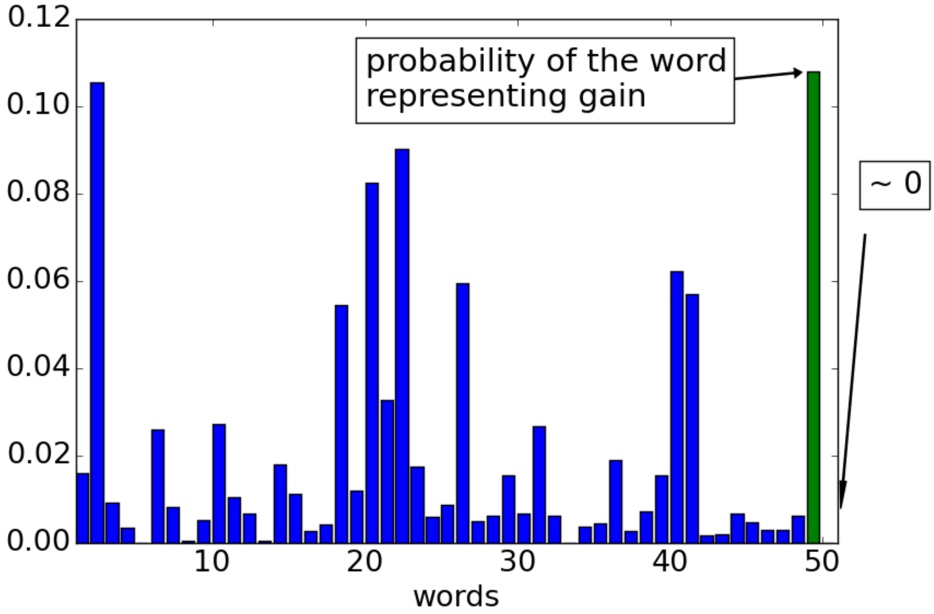
$$\boldsymbol{w} = (\text{"ag2"}, \text{"ag3"}, \text{"ag3"}, \text{"loss"}, \text{"loss"}, \text{"loss"}, \text{"loss"})$$

Since LDA is a bag-of-words model, the order of the words in the document does not matter. The batch of documents the online LDA model of agent i is supplied with at iteration t, consists of the interpreted-as-documents coalitions i has joined at t.

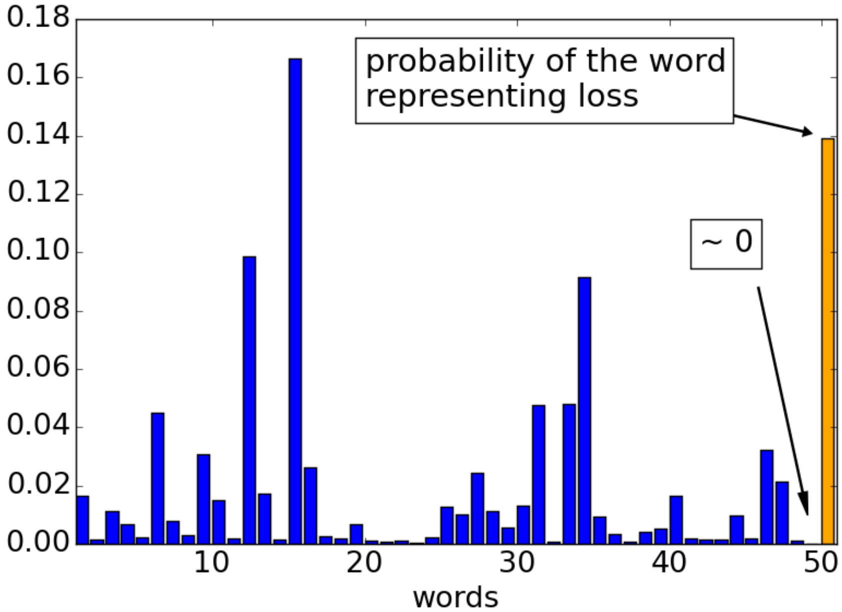
The intuition behind the notion of a topic is that the words that appear in it with high probability are very likely to appear together in a document that exhibits this topic with high probability. The probability with which the word corresponding to an agent's contribution appears in a topic is correlated with the amount of her contribution. Therefore, the meaning of a topic identified by agent i, is that i has observed in many documents certain agents who contributed a lot, and some that contributed less; and this configuration results to gain or loss with the corresponding probabilities.

Thus, the topic in Fig. 6.1(a) implies that if i joins a coalition with the agents that appear in the topic with high probability, then that coalition would be profitable. On the other hand, the topic in Fig. 6.1(b) implies that forming a coalition with the agents that appear in it with high probability would result in loss. Note that learning a topic's profitability corresponds to acquiring information on the RRs associated with that topic. However, these RRs are not explicitly learned; what is learned is the underlying collaboration structure (which might, in the general case, be generated by means other than RRs). It is natural to expect that agents that appear with high probability in a topic which has been associated with loss, like the one in Fig. 6.1(b), will not appear (as a group) with high probability in a topic that has been associated with gain, like the one in Fig. 6.1(a). Such occurrence would reflect that an agent's beliefs indicate that cooperation with a group of agents is (paradoxically) both beneficial and harmful. Furthermore, online LDA diminishes the problem caused by an agent's lack of knowledge of the resource quantity that others hold, since topic updates are based on observations of words. Therefore, if an agent possesses a small resource quantity, the word that corresponds to her contribution will probably be observed fewer times than those of the other agents, and it will thus affect the formation of the topics analogously.

[2]The number of times the word for utility is written may require scaling in cases where its domain ranges from very low to values with a much higher order of magnitude.



(A) A "profitable" learned topic.



(B) A "non-profitable" learned topic.

FIGURE 6.1: Typical topics, as formed by a randomly selected agent at the end of a random iteration in an experiment, where an agent's vocabulary consists of 51 words ($n = 50$). The two last words in a topic indicate the probability of gain and loss respectively, while the rest correspond to agents' contribution. In (a), the "profitable" learned topic, the word for loss appears with near-zero probability; in (b), the "non-profitable" topic, the word for gain has near-zero probability.

6.3 Taking Formation Decisions in Repeated OCF Games

In this section we present our method, which is used for agent decision-making. It consists of two algorithms; one used by the proposer of an iteration in order to propose coalitions (and offer them some resource amount), and another which is used by the proposed agents in order to respond to the received proposals, i.e., accept (and offer some resource quantity) or reject. We term it `OVERPRO`, since it has been developed for OVERlapping coalition formation via PRObabilistic topic modeling (employing online LDA). `OVERPRO` is used assuming a simple formation protocol, but this can be replaced by any protocol of one's choosing.

6.3.1 A Repeated OCF Protocol

We now proceed to present this protocol. The protocol of our game operates in I iterations (rounds). At the beginning of an iteration one agent is randomly selected, from the set of agents $N = \{1, \ldots, n\}$, as the proposer, and is thus given the ability to propose a number of partial coalitions, offering an integer quantity of her resource to each of them. Therefore, proposer $i \in N$ is asked to pass a list of tuples of the form $\langle C, r_{i,C} \rangle$, where $C \subseteq N$ and $r_{i,C} \in \mathbb{N}^+$ denotes the amount of resource that i offers to coalition C. By limiting the offers of i to discrete quantities we disallow the proposal of an infinite number of coalitions. Then, every agent $j \in N \setminus i$ is a responder and gets informed of the proposals in which she is involved—i.e., of the proposals $\langle C, r_{i,C} \rangle : j \in C$, and decides on the quantity $r_{j,C} \in \mathbb{N}$ to invest in C, where $r_{j,C} = 0$ essentially means that j has declined participation in C. The total resource that j offers to coalitions cannot exceed r_j (the same holds for the proposer). An agent makes her decisions without communication with any other agent. A (partial) coalition C forms if and only if all involved agents accept to participate in it.

The resources of the agents are *replenished* at the end of each iteration, and hence there is no need for strategic planning, on the part of the agents, which would have necessitated had the resources not be renewed. Moreover, at the end of each round all coalitions are dissolved. The utility $u_{i,C}$ that $i \in C$ earns from coalition C is *proportional* to her $r_{i,C}$ contribution, i.e., $u_{i,C} = u_C \cdot r_{i,C} / \sum_{j \in C} r_{j,C}$, where u_C is the total utility earned by the coalition. An agent receives contributions and utility information only about her formed coalitions.

6.3.2 The OVERPRO Method

We now present `OVERPRO`, our novel method that exploits probabilistic topic modeling for decision-making and which can be easily adjusted to other protocols. The main idea behind `OVERPRO` is that an agent, by updating her online LDA model throughout the game, forms topics which indicate how well she can cooperate with others. Thus, by considering as "profitable" ("non-profitable") the topics in which the probability of the word that represents gain (loss) is higher than the probability of the one that represents loss (gain), the agent can identify coalitions that will potentially result in gain (loss)—positive utility (negative utility). However, it might be that not all of the topics are *significant*, since some of them might not be well formed (especially at the early iterations of the game). An agent should not make decisions based on non-significant topics, since they may provide inaccurate information and lead to harmful actions, or the prevention of making beneficial ones. Thus, a topic (out of the K topics of the agent's model) will be either *significant* or *non-significant*. We define a topic to be *significant* if the absolute value of the difference between the probability of the word representing gain and the probability of the word representing loss is greater than ϵ. Agent i's significant topics are denoted as ST^i:

$$ST^i \leftarrow \{k : |\beta^i_{k,'gain'} - \beta^i_{k,'loss'}| > \epsilon\}$$

where β^i_k is the k^{th} (out of the K) topic of agent i. We further define *Good* (profitable) topics as the ones in which the probability of the word representing gain is greater than that of the word representing loss, and are significant. *Bad* topics are analogously defined. Naturally, we use the following notation, for the *Good* and *Bad* topics of agent i:

$$Good^i \leftarrow \{k : \beta^i_{k,'gain'} > \beta^i_{k,'loss'} \wedge k \in ST^i\}$$
$$Bad^i \leftarrow \{k : \beta^i_{k,'gain'} < \beta^i_{k,'loss'} \wedge k \in ST^i\}$$

Moreover, it must be noted that every word in a topic appears with positive (no matter how small) probability, due to the initial randomization of λ which is the (variational) hyperparameter of the topics. (The probability of an event with Dirichlet prior cannot be zero, otherwise there would not be a corresponding parameter in the Dirichlet.) However, as an agent learns, some of them tend to zero. But still, How much probability should an agent appear with in a topic in order to be considered significant? For example, in Fig. 6.1(a) not all agents appear in the profitable topic with similar probability values. We define the *significant agents* of topic k of agent i, denoted as SA^i_k, as those whose corresponding words in topic k have probability higher than the mean value μ of the probabilities of the words corresponding to agents, plus the standard deviation σ of those. Formally, the set SA_k^i of significant agents of topic k as identified by agent i, is defined as follows:

$$SA_k^i \leftarrow \{j : \beta_{k,j}^i > \mu(\beta_{k,N\setminus i}^i) + \sigma(\beta_{k,N\setminus i}^i) \wedge j \in N \setminus i\}$$

The proposer i at iteration t has to make proposals of the form $\langle C, r_{i,C} \rangle$, $C \subseteq N$. Therefore, i must not only decide on which coalitions to propose to, but on how much of her resource r_i to invest to each of them also. The approach of OVERPRO is to propose the coalitions that are indicated by the profitable topics. The resource quantity offered to a coalition is proportional to the quality of the topic as modelled by i, where the quality of topic k is defined as the difference between the probabilities of the words indicating gain and loss—i.e., $qual_k^i = \beta_{k,'gain'}^i - beta_{k,'loss'}^i$. Furthermore, the quantity $r_{i,C}$ offered to C must be integer-valued. Thus, $r_{i,C}$ has to be rounded appropriately—i.e., the finally offered value will be either $\lfloor r_i, C \rfloor$ or $\lfloor r_{i,C} \rfloor + 1$, since the total offered quantity must not exceed the resource of i. Picking just the most profitable topic, and thus proposing a single coalition, is highly risky, since the rejection by one responder is sufficient to not allow its formation. Furthermore, it is highly likely that at the early iterations of a game the beliefs of the agents, represented by topics, are not accurate. Thus, agents propose multiple (overlapping) coalitions.

A notorious problem in learning is the exploitation-vs-exploration (Sutton and Barto, 1998) one: an agent must choose to either exploit her best-so-far action, or explore different options. We deal with this issue by allowing agent i to do both at the same time, since r_i is divisible. Specifically, at iteration t proposer i dedicates $z_t \in (0, 1)$ of r_i in exploring and $1 - z_t$ in exploiting. Then, an agent performs exploration by proposing $\lfloor r_i \cdot z_t \rfloor$ coalitions, offering to each the minimum possible resource quantity $(= 1)$. Alg. 7 below depicts the proposer's decision-making (with Alg. 8 describing her exploration process).

In each iteration, responders receive the proposals in which they are involved, and decide, for each proposal in turn whether to accept it (invest to it some positive resource quantity) or reject it (offering nothing). OVERPRO employs a parameter $c \in (0, 1)$, so that an agent rejects a proposed coalition C if she identifies a non-profitable (bad) topic in which at least $(c \cdot 100)\%$ of the agents in C are significant. The intuition behind the employment of parameter c is that it suffices to observe a certain percentage of agents of a proposed coalition in a "non-profitable" topic in order to reject it. Parameter c can have different values at different rounds, so we refer to its value at iteration t as c_t.

If a coalition is not rejected, then it is checked whether there is a profitable (good) topic in which at least $((1 - c_t) \cdot 100)\%$ of the agents in C are significant. In that case the resource invested in C is proportional to its quality, otherwise a minimum quantity is offered. The intuition is that an agent can classify a coalition C in three categories: It can be the case that C is bad, as indicated by the topics, and thus nothing should be offered. Topics may provide no

Algorithm 7: OVERPRO-propose (by proposer i)

$amount\_explore \leftarrow \lfloor r_i \cdot z_t \rfloor$
$amount\_exploit \leftarrow r_i - amount\_explore$
$ST^i \leftarrow \{k : |\beta^i_{k,'gain'} - \beta^i_{k,'loss'}| > \epsilon\}$
$Good^i \leftarrow \{k : \beta^i_{k,'gain'} > \beta^i_{k,'loss'} \wedge k \in ST^i\}$
$\forall k \in Good^i \ qual^i_k \leftarrow \beta^i_{k,'gain'} - \beta^i_{k,'loss'}$, $SA^i_k \leftarrow \{j : \beta^i_{k,j} > \mu(\beta^i_{k,N\setminus i}) + \sigma(\beta^i_{k,N\setminus i}) \wedge j \in N \setminus i\}$
$Proposals \leftarrow \emptyset$
for $k \in Good^i$ **do**
    $C \leftarrow SA^i_k \cup i$
    $r_{i,C} \leftarrow amount\_exploit \cdot qual^i_k / \sum qual^i$
    round $r_{i,C}$ appropriately-do not exceed $amount\_exploit$
    $Proposals.add(\langle C, r_{i,C} \rangle)$
return $Proposals \ \cup \ explore(amount\_explore)$

Algorithm 8: explore

$Proposals \leftarrow \emptyset$
for $\{1, \ldots, amount\_explore\}$ **do**
    $size \leftarrow$ random($\{1, \ldots, n-1\}$)
    $C \leftarrow$ (randomly choose $size$ agents from $N \setminus i) \cup i$
    $Proposals.add(\langle C, 1 \rangle)$
return $Proposals$

Algorithm 9: OVERPRO-respond (by responder i)

$ST^i \leftarrow \{k : |\beta^i_{k,'gain'} - \beta^i_{k,'loss'}| > \epsilon\}$
$Good^i \leftarrow \{k : \beta^i_{k,'gain'} > \beta^i_{k,'loss'} \wedge k \in ST^i\}$
$Bad^i \leftarrow \{k : \beta^i_{k,'gain'} < \beta^i_{k,'loss'} \wedge k \in ST^i\}$
$\forall k \in ST^i \ SA^i_k \leftarrow \{j : \beta^i_{k,j} > \mu(\beta^i_{k,N\setminus i}) + \sigma(\beta^i_{k,N\setminus i}) \wedge j \in \ N \setminus i\}$
$Responses \leftarrow \emptyset$; $Rejected \leftarrow \emptyset$
for $C \in CoalitionsOfReceivedProposals$ **do**
    for $k \in Bad^i$ **do**
        if $|SA^i_k \cap C| > c_t \cdot |C|$ **then**
            $Responses.add(\langle C, 0 \rangle)$; $Rejected.add(C)$
for $C' \in CoalitionsOfReceivedProposals \setminus Rejected$ **do**
    $Responses.add(\langle C', 1 \rangle)$
    $r_i \leftarrow r_i - 1$
$\forall k \in Good^i \ qual^i_k \leftarrow \beta^i_{k,'gain'} - \beta^i_{k,'loss'}$
Assign to each coalition C' that has not been rejected the quality of the (good) topic with the highest one among those (if any) which satisfy $|SA_k \cap C'| > (1 - c_t) \cdot |C'|$; distribute the remaining r_i proportionally to these qualities; update $Responses$ since resource offerings have increased
return $Responses$

particular information regarding C and thus a minimum quantity can be offered, so that experience from participation is gained. If the case is that C is indicated as beneficial, then an additional investment should be offered, with the exact offering depending on how beneficial C is compared to the other good ones. The entire response-formation process is shown at Alg. 9.

The training of LDA, and thus consequently that of online LDA also, takes polynomial time (Blei, Ng, and Jordan, 2003), as a result of the variational inference. Furthermore, as it can be observed in Alg. 7, the number of computation steps of `OVERPRO`-propose is polynomial in the number of topics. From Alg. 9, we can see that the running time of `OVERPRO`-respond is polynomial in the number of coalitions to respond to. Therefore, the number of computation steps that are involved in `OVERPRO` is polynomial. In the exploration phase any algorithm of one's choosing can be used, as it is independent to `OVERPRO`. Despite that the one presented in Alg. 8 runs in pseudopolynomial time, as it depends in the quantity of an agent which is a numeric value, it is natural to assume that agents' resource quantities are not large values, and thus this does not cause any problems in practice.

6.4 Q-Learning for Overlapping Coalition Formation

To the best of our knowledge, this is the first work on (decentralized) multi-agent learning for *overlapping* coalition formation under uncertainty, and thus there is no algorithm to use as a means for comparison. To meet this need, we have developed a Q-learning-style (Watkins and Dayan, 1992; Claus and Boutilier, 1998) algorithm as a baseline. An agent that uses our Q-learning algorithm employs two distinct kinds of Q-values. The first one, which is denoted as Q_a, maintains agent-level values; while the second, which is denoted as Q_s, maintains coalition size-level values. Employing two different sets of Q-values is necessary since the alternative of maintaining a Q-value for every possible coalition requires exponential space in the number of the agents (rendering the problem practically intractable in large settings). Agent i maintains for each agent $j \in C \backslash i$ a $Q^i_{a,j}$ value, and for each $m \in \{1, \ldots, n-1\}$ a $Q^i_{s,m}$ value; keeping a $Q^i_{s,m}$ value for $m = n$ is redundant since the decision-maker always includes herself in a coalition. Furthermore, a learning rate $\delta_t \in (0, 1)$ is employed (Sutton and Barto, 1998), as is common in Q-learning, where t is the game iteration. After $C, i \in C$, is formed, agent i updates her Q-values as follows:

$$Q^i_{a,j} \leftarrow Q^i_{a,j} + \delta_t \left((u_C / r_{-i,C}) - Q^i_{a,j} \right) \ \forall j \in C \setminus i$$
$$Q^i_{s,m} \leftarrow Q^i_{s,m} + \delta_t \left((u_C / r_{-i,C}) - Q^i_{s,m} \right) \text{ where } m = |C| - 1$$

where $r_{-i,C}$ is total resource investment in C excluding that of i. Thus, Q-values correspond to efficiency indices, i.e. the ratio of utility u_C to the total resource quantity invested in C.

A proposer employing our Q-learning algorithm iteratively selects some quantity of her resource to offer to a coalition, until it is depleted. Then, the size of the coalition to propose (excluding herself) is selected using the *softmax function* (Sutton and Barto, 1998) over the Q^i_s values, and afterwards the agents to include in the coalition are selected using the *softmax function* over the Q^i_a values. The approach to the exploitation-vs-exploration problem is exactly the same as in `OVERPRO`, and thus the z_t parameter is also employed here, as well. Alg. 10 depicts the proposer's decision-making process.

Algorithm 10: Q-learning-propose (by proposer i)

$amount\_explore \leftarrow \lfloor r_i \cdot z_t \rfloor$
$amount\_exploit \leftarrow r_i - amount\_explore$
$Proposals \leftarrow \emptyset$
while $amount\_exploit > 0$ **do**
  $quantity \leftarrow$ random$(\{1, \ldots, amount_exploit\})$
  $amount\_exploit \leftarrow amount\_exploit - quantity$
  $size \leftarrow$ select coalition size with probability $e^{Q^i_{s,m}} / \sum_{m'} e^{Q^i_{s,m'}}$, where $m \in \{1, \ldots, n-1\}$
  $C \leftarrow$ (select $size$ agents with probability $e^{Q^i_{a,j}} / \sum_{j'} e^{Q^i_{a,j'}}$, where $j \in N \setminus i$) $\cup\, i$
  $Proposals.add(\langle C, quantity \rangle)$
return $Proposals \cup explore(amount\_explore)$

Agent i accepts participation in a proposed coalition C with certainty if $\sum_{j \in C \setminus i} Q^i_{a,j}$ is positive, or else the proposal is accepted with probability $h_t \in (0, 1)$. A minimum quantity is initially offered to every accepted coalition, and the remaining quantity is distributed to the coalitions that are accepted with certainty, proportionally to the sum of their Q-values. The exact way agents use Q-values to respond is shown in Alg. 11.

6.5 Experiments

We evaluated `OVERPRO`'s effectiveness and robustness in environments with 50 and 250 agents. Agent resource quantities were generated from $\{450, \ldots, 550\}$ uniformly at random. The RRs were 400 for $n = 50$, and 20k for $n = 250$. Every game ran for I = 1000 iterations, and thus, agent i can observe at most $r_i \cdot$ 1000 documents. The development is in Python 3 and Online LDA was implemented as in (Hoffman, Bach, and Blei, 2010).[3] The same exploration rate z_t was set for both `OVERPRO` and Q-learning, decreasing quadratically

[3] https://github.com/blei-lab/onlineldavb

Algorithm 11: Q-learning-respond (by responder i)

$Responses \leftarrow \emptyset$; $Beneficial \leftarrow \emptyset$
for C *in* $CoalitionsOfReceivedProposals$ **do**
  $qual^i_C \leftarrow \sum_{j \in C \setminus i} Q^i_{a,j}$
  if $qual^i_C > 0$ **then**
    $Responses.add(\langle C, 1 \rangle)$
    $r_i \leftarrow r_i - 1$
    $Beneficial.add(C)$
  else if $random(0,1) < h_t$ **then**
    $Responses.add(\langle C, 1 \rangle)$
    $r_i \leftarrow r_i - 1$
  else
    $Responses.add(\langle C, 0 \rangle)$
Distribute the remaining resource r_i to coalitions in $Beneficial$ proportionally to their $qual$ values; update $Responses$ since resource offerings have increased
return $Responses$

from 1 to 0.01. Both c_t and h_t decreased linearly from 1 to 0.35. The exact values of z_t, c_t, and h_t over time are presented in the Appendix. We tested both OVERPRO, which requires K (number of topics),[4] τ_0 and κ (that determine the impact $\rho = (\tau_0 + t)^{-\kappa}$ of a batch of documents on the topics), and Q-learning, which requires δ_t, for a number of different parameters. Experiments were conducted on three i3 3GHz PCs with 4GB of RAM and two cores each; on average a game for $n = 50$ took about 6,500 sec in total, and 95,000 sec for $n = 250$.

In Table 6.1 we present for $n = 50$ the average: social welfare earned in a game (sw); number of coalitions that a proposer of an iteration proposes (proposals); number of coalitions in which an agent participates in a round (participation); and time it took to complete a game per iteration per agent.

As observed in Table 6.1, OVERPRO performs much better than Q-learning in terms of social welfare. For the best set of parameters of OVERPRO, $\langle K = 10, \tau_0 = 200, \kappa = 0.9 \rangle$, the average social welfare earned in a game was about the triple of that earned when Q-learning with the best value $\delta_t = 0.95^t$ was employed, since $1320.24/475.31 = 2.77$. For Q-learning, the performance for $\delta_t = 0.995^t$ was less than the half of the one for $\delta_t = 0.95^t$, as $205.73/475.31 = 0.43$. OVERPRO performed better when the number of topics K maintained by every agent was 10, rather than 15, for both pairs of $\langle \tau_0 = 200, \kappa = 0.9 \rangle$ and $\langle \tau_0 = 100, \kappa = 0.7 \rangle$. For both 10 and 15 topics, the social welfare was better for $\langle \tau_0 = 200, \kappa = 0.9 \rangle$ than for $\langle \tau_0 = 100, \kappa = 0.7 \rangle$. Since τ_0 and κ determine the impact that a batch of documents has on the formation of the

[4] In some LDA implementations the value of K is automatically derived Teh et al., 2006, but we exploit the standard online LDA algorithm, which requires passing the value of K as parameter.

TABLE 6.1: Results (averages over 100 runs) for 50 agents and different values of $\langle K, \tau_0, \kappa\rangle$ for OVERPRO and δ_t for Q-learning. Proposals, participation and time (in sec) are per agent per iteration (there is a unique proposer in an iteration).

$n = 50$	sw ($\cdot 10^3$)	proposals	participation	time (sec)
$\langle 10, 200, 0.9\rangle$	1320.24	111.28	49.68	0.3808
$\langle 10, 100, 0.7\rangle$	1183.94	110.63	49.48	0.3479
$\langle 15, 200, 0.9\rangle$	1237.42	113.13	49.73	0.4074
$\langle 15, 100, 0.7\rangle$	982.40	112.20	49.39	0.3691
$\delta_t = 0.95^t$	475.31	113.81	47.86	0.0054
$\delta_t = 0.9^t$	286.08	113.80	37.81	0.0079
$\delta_t = 0.995^t$	205.73	113.81	27.35	0.0044

topics, $\rho_t = (\tau_0{+}t)^{-\kappa}$ can be interpreted as a learning rate. Now, higher values of τ_0 and κ result in smaller values of ρ_t. Therefore, it can be conjectured that slower learning rates are preferred over faster ones.

The number of proposals made in an iteration is largely affected by the exploration phase. Since we have experimented with high values of exploration rate (the average value of z_t is 0.215), the number of proposals is almost the same for all values of $\langle K, \tau_0, \kappa\rangle$ for OVERPRO. However, we notice that a proposer employing OVERPRO with higher values of τ_0 and κ tends to suggest a greater number of coalitions than one with smaller values of τ_0 and κ does, since for both 10 and 15 topics the number of proposals is greater for $\langle \tau_0 = 200, \kappa = 0.9\rangle$ than for $\langle \tau_0 = 100, \kappa = 0.7\rangle$, as $111.28 > 110.63$ and $113.13 > 112.20$. Agents employing OVERPRO, in general tend to join more coalitions than their Q-learning counterparts, as indicated by the values of "participation" in Table 6.1, which are greater for the former. By the end of a game an agent employing OVERPRO will have trained her online LDA with more than 49k documents, since one coalition corresponds to one document, an agent participates in at least 49.39 coalitions in a round $\big($value for $\langle K = 10, \tau_0 = 100, \kappa = 0.7\rangle\big)$, and I=1000.

Now, one cannot draw accurate conclusions regarding the real power of an agent decision-making algorithm relying solely on social welfare. For instance, if one makes many proposals, it is more likely that more agents will join, and thus the social welfare will probably increase. Therefore, we define *efficiency* as the ratio of social welfare (total utility) to total resource quantity invested by all agents in every coalition in a round.[5] This efficiency metric is quite natural, since it takes the focus away from social welfare—and rational agents aim to maximize their own utility, and not the social welfare.

It can be observed in Fig. 6.2(a) that OVERPRO vastly outperforms the Q-learning algorithm in terms of efficiency. The efficiency of Q-learning hardly

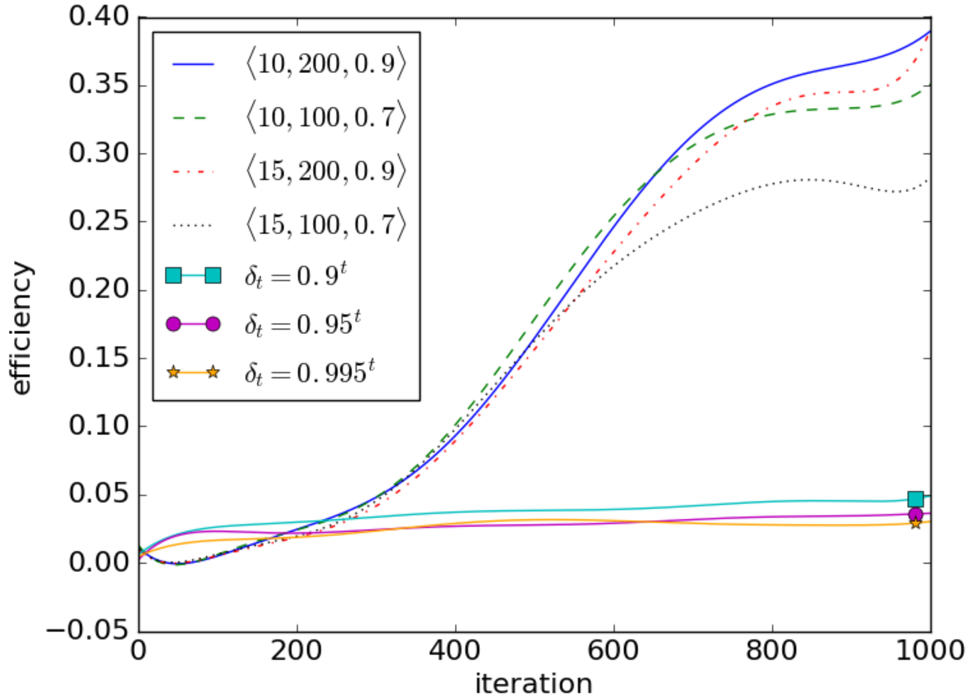
[5] One problem in plotting efficiency is that since the dynamics of the system are complex (as a result of the form of the protocol and the number of the agents), there are multiple fluctuations and overlaps that make the plots hard to make out. Thus, we employed curve fitting with a polynomial function.

increases throughout the game, whereas the performance of `OVERPRO` improves steadily. We can thus conclude that as iterations go by, agents employing `OVERPRO` are more efficient in terms of earning utility (as a function of resources invested). In the Appendix, we present the number of formed coalitions for each set of experiments over time, and the (average) perplexity of an agent's online LDA model for `OVERPRO`.

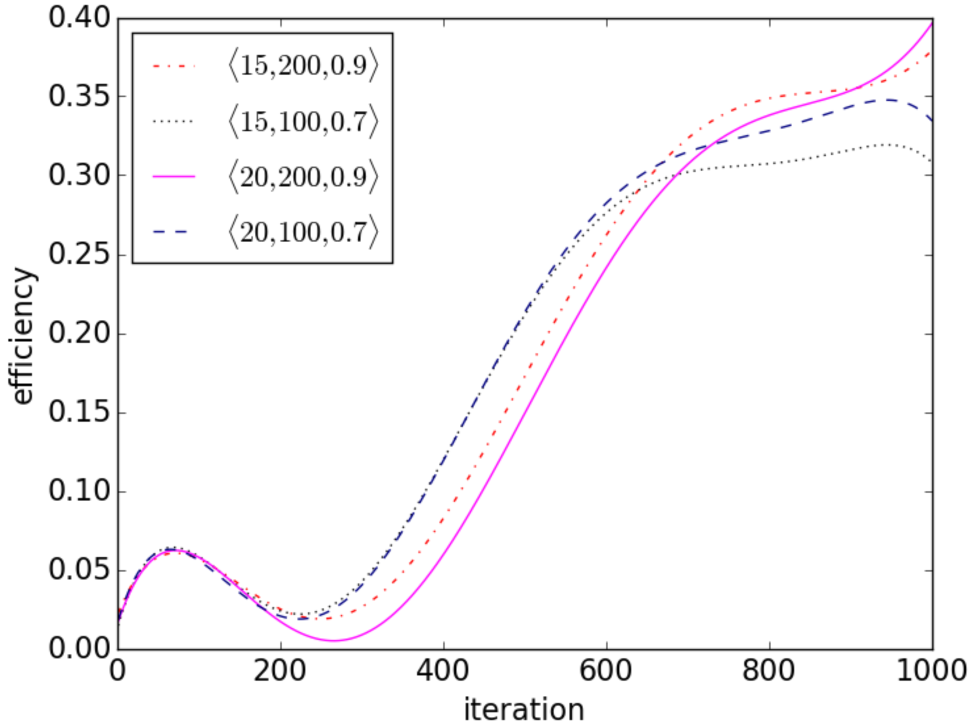
Now, as depicted in Fig. 6.2(b), `OVERPRO` achieves even better efficiency for $n = 250$ than for $n = 50$, while its behaviour through time is very similar. Therefore, it exhibits robust performance against increases of n. Furthermore, we have observed through experimentation that the number of topics K should increase sublinearly to n. Agents using Q-learning, for $n = 250$, were for most rounds unable to form coalitions, as they could not coordinate their actions, and thus the corresponding efficiency could not be plotted. This is a problem that Q-learning algorithms face when they are employed in large-scale systems (Agogino and Tumer, 2005). However, agents using `OVERPRO` overpass this problem. In fact, each agent had trained her online LDA model with more than 18.5k documents (coalitions) in total, at the end of the game. The decrease of the number of coalitions that an agent is member of in a game with $n = 250$ compared to $n = 50$ is natural, since the protocol in effect demands that all of the proposed agents of a potential coalition agree on its formation without allowing any further interaction. Nevertheless, even for $n = 250$ agents using `OVERPRO` achieve to form of a fairly large number of coalitions.

6.6 Conclusions

We have presented a novel approach towards multi-agent learning in cooperative game environments, where probabilistic topic modeling, and specifically online LDA, is exploited. Furthermore, this is the first work to tackle overlapping coalition formation under uncertainty, where the uncertainty is on the relations entailing synergies among the agents. To this end, we proposed *Relational Rules* (RRs), a representation scheme which extends MC-nets to cooperative games with overlapping coalitions; and then showed how to use online LDA to implicitly learn the agents' synergies described by (unknown) RRs. Our method, `OVERPRO`, decisively outperforms a Q-learning algorithm we also developed.



(A) $n = 50$. Results are averages over 70 runs.



(B) $n = 250$. Results are averages over 30 runs.

FIGURE 6.2: Efficiency defined as the ratio of social welfare (total utility) to total resource quantity invested by all agents in every coalition in an iteration (round) for different values of $\langle K, \tau_0, \kappa \rangle$ for OVERPRO and δ_t for Q-learning (for $n = 50$).

Chapter 7

Conclusions

This thesis provides the first approach in the literature for overlapping coalition formation under uncertainty. We have presented and combined approaches that had never been taken under consideration before in the coalition formation and cooperative game theory literature. In particular, the main concepts we have exploited come from the fields of Probabilistic Computing, and especially Probability Inequalities, and Probabilistic Topic Modeling. Our experimental results demonstrate the effectiveness of our approach, and its potential for practical usage in real world environments. In the following subsections, we provide a summary of our approaches and contributions, and then discuss topics and extensions that could be studied in future work.

7.1 Summary

In the first of our works, presented in section 5, we have exploited Probability Bounds, given by Probability Inequalities, for Overlapping Coalition Formation. We considered a model where agents have uncertainty regarding the resource quantity that the others possess, and provided a way for modeling an agent's beliefs for the potential resource investment of the others. In this, context we used Beta and Dirichlet conjugate priors, providing thus a Bayesian approach for modeling the uncertainty of the agents. Thereafter, we provided three methods, exploiting an improvement of the Paley-Zygmund inequalty, the two-sided Chebyshev's inequality, and the Hoeffding's inequality. These methods allow agents to demand certain confidence levels and take formation decisions by approximating the distributions derived by their beliefs, as their exact computation is excessively time consuming, and taking advantage of the derived bounds. In order to provide a means for the agents to form coalitions, we have presented an overlapping coalition formation protocol, under which proposers make offers (promising a portion of the target utility) to the rest of the agents towards task completion, which results in acquiring utility.

We have conducted experimental evaluation of our methods by testing them on both a random graph of 300 nodes (agents) and snapshot of Facebook consisting of 4039 nodes. We used as a baseline a method that takes decisions based on expected resource contribution. Our methods consistently outperformed the baseline, in terms of task completion, and exhibited similar behaviour in the two graphs.

In our second work, we exploited Probabilistic Topic Modeling in order to devise an effective method for agent decision-making in iterated overlapping coalition formation environments, where agents have uncertainty regarding the underlying collaboration structure. In particular, we took advantage of online Latent Dirichlet Allocation and showed how an agent can learn profitable coalitions by interpreting coalitions as documents. Furthermore, we have proposed Relational Rules, a novel representation scheme for cooperative games with overlapping coalitions which extends the well-known MC-nets presentation.

We have also presented `OVERPRO`, a method that exploits beliefs as shaped by online Latent Dirichlet Allocation, and compared it to a Q-learning style algorithm that we devised, as well. `OVERPRO` exhibited vastly better performance compared to the Q-learning algorithm. Therefore, we can conclude that the underlying collaboration structure, described by Relational Rules, is efficiently learned and exploited by `OVERPRO`.

7.2 Future Work

Here we outline our ongoing and future work. Both pillars of this thesis allow for extensions that can possibly lead to even better results or new research directions.

First, we consider testing our proposed methods for deriving probability bounds on different coalition formation protocols so as to confirm their robustness. Additionally, we intend to conduct experiments where agent resources will be depleted, instead or renewed, over time. Naturally, this would necessitate long-term strategic planning. It is expected that such a setting would have an impact on the way that agents derive bounds over uncertainty.

Our work on overlapping coalition formation via probabilistic topic modeling essentially presents a brand-new approach to multi-agent learning in cooperative games. Thus, it is desirable to formulate our proposed method so that we employ it in non-cooperative environments. Moreover, we intend to study how scaling the number of times that words are written in a document affects the formation of the topics, and thus the effectiveness of our method. Furthermore, we intend to explore alternative approaches for exploiting the knowledge provided by the learned topics.

Finally, we are interested in intertwining probability bounds and probabilistic topic modeling. Such a combination could lead to even more interesting research directions and results. It is intriguing to study the way that such an approach in coalition formation settings would be related to concepts from non-cooperative game theory.

Appendix A

Appendix

In this appendix we present figures that were omitted in Chapter 6.

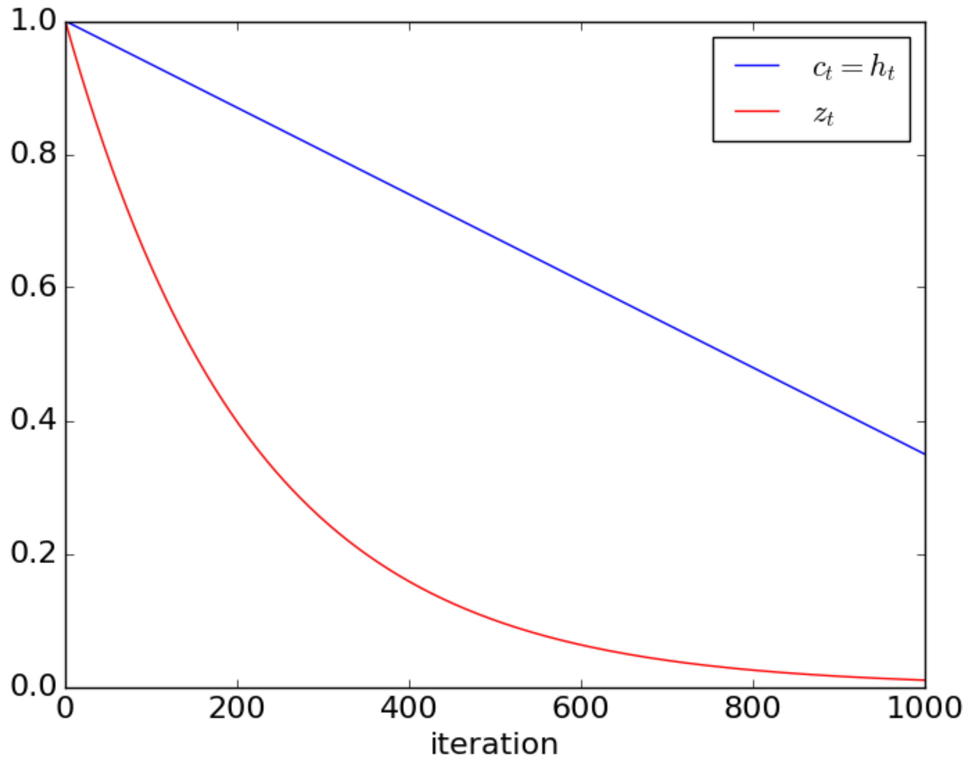
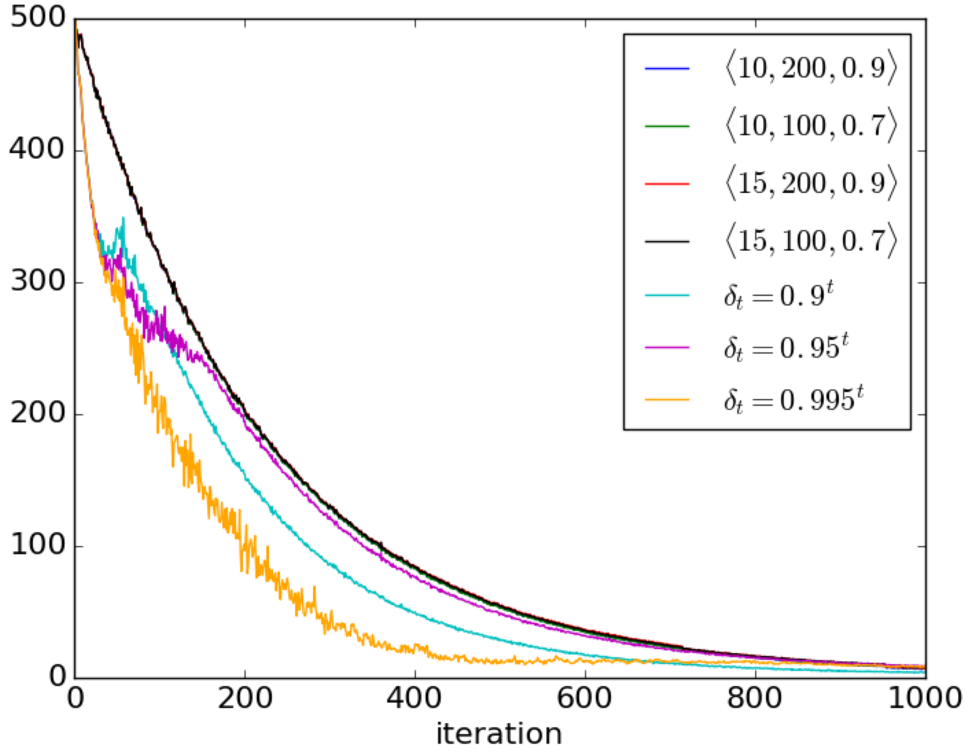


FIGURE A.1: Exploration rate z_t, and parameters c_t and h_t that affect the responses of the agents.

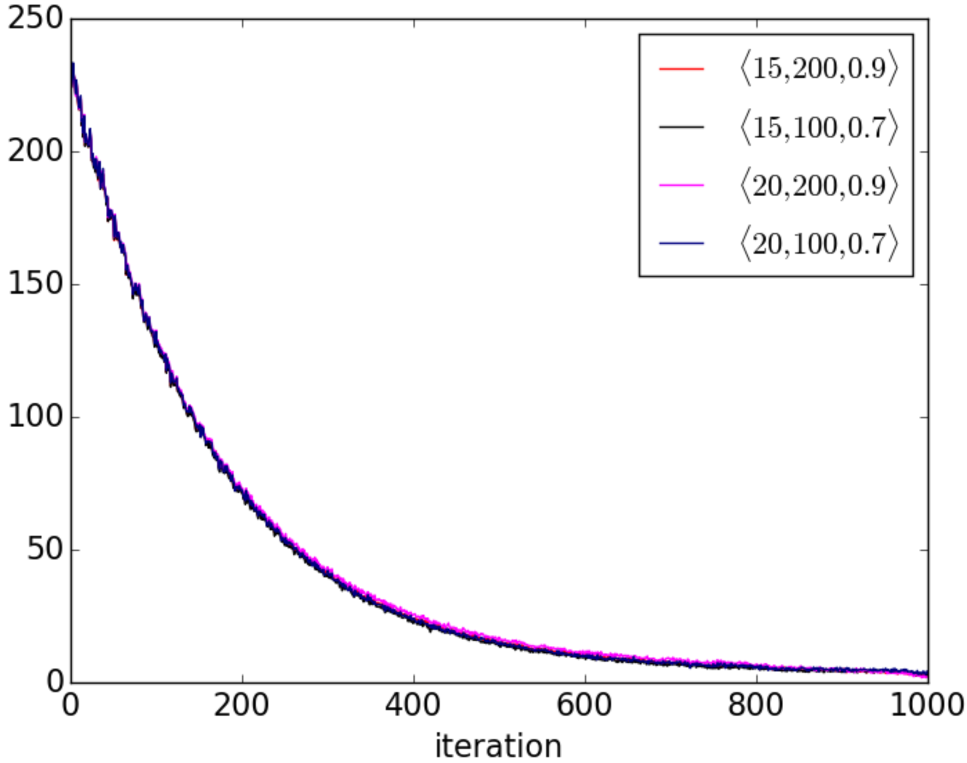
In Figure A.1 we observe the exploration rate z_t that was used by both OVERPRO and Q-learning in all of the experiments. In the first iteration its value is 1 and it finally reaches 0.01 by decreasing in a quadratic fashion. Both parameters c_t, used by OVERPRO, and h_t, used by the Q-learning algorithm, are initially valued at 1, and linearly decrease to 0.35. As stated in Chapter 6, c_t and h_t are equally valued since they affect the decisions of the responders in a similar manner.

In Figures A.2 (A) and A.2 (B) we observe the number of coalitions that form in each iteration, for $n = 50$ and $n = 250$, respectively. For $n = 50$, we can see that number of formed coalitions in every iteration is almost the same for all configurations of OVERPRO (four first labels) and Q-learning for $\delta_t = 0.95^t$. Furthermore, we observe that for Q-learning with $\delta_t = 0.995^t$ the number of formed coalitions is much lower than for any other setting until the 700^{th} iteration, where from there on Q-learning for $\delta_t = 0.9^t$ is the setting for which the least number of coalitions per iterations forms. For $n = 250$, and for every set of $\langle K, \tau_0, \kappa \rangle$ of OVERPRO, the number of formed coalitions is virtually the same. We can see that, in general, the rate of decrease of the number of formed coalitions for $n = 250$ is higher than the one for $n = 50$.

In figures A.3(A) and A.3(B) we see the average values of perplexity of the online LDA models of the agents over time, for $n = 50$ and $n = 250$, respectively. Naturally, in either case, in the first iteration perplexity has a high value, since the variables involved in the online LDA models have been randomly initialized. For $n = 50$, we see that the lowest perplexity is observed for $tau_0 = 200$ and $\kappa = 0.9$, and there is virtually no difference for either $K = 10$ or $K = 15$. However, no significant difference is observed for different values of $\langle K, \tau_0, \kappa \rangle$ for $n = 250$.

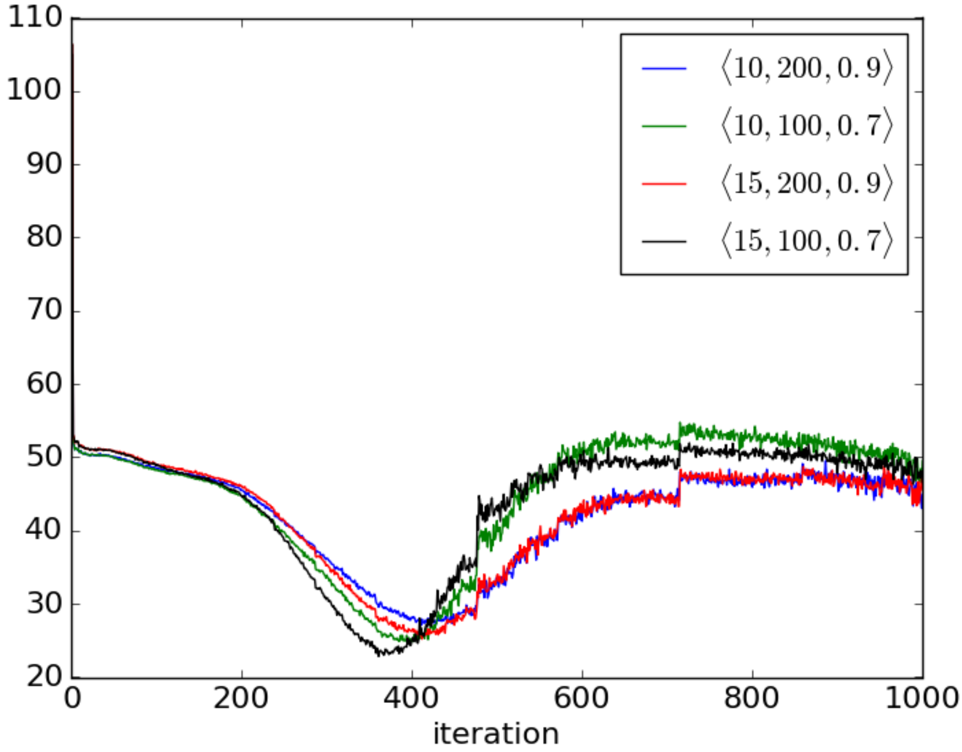


(A) $n = 50$

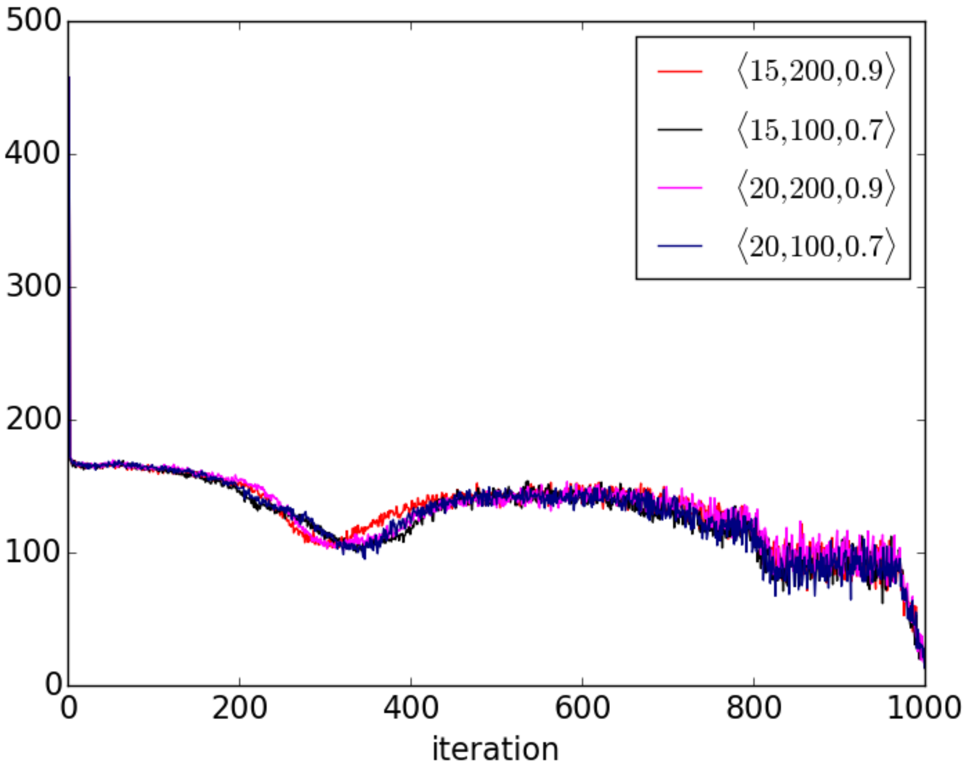


(B) $n = 250$

FIGURE A.2: Number of formed coalitions per iteration.



(A) $n = 50$



(B) $n = 250$

FIGURE A.3: Perplexity per iteration.

Bibliography

Aadithya, Karthik V, Tomasz P Michalak, and Nicholas R Jennings (2011). "Representation of coalitional games with algebraic decision diagrams". In: *Proc. of the 10th International Conference on Autonomous Agents and Multiagent Systems*, pp. 1121–1122.

Abdi, Hervé and Lynne J Williams (2010). "Principal component analysis". In: *Wiley interdisciplinary reviews: computational statistics* 2.4, pp. 433–459.

Agogino, Adrian and Kagan Tumer (2005). "Reinforcement learning in large multi-agent systems". In: *AAMAS-05 Workshop on Coordination of Large Scale Multi-Agent Systems*.

Apt, Krzysztof R, Sunil Simon, and Dominik Wojtczak (2015). "Coordination games on directed graphs". In: *Proc. of 15th International Conference on Theoretical Aspects of Rationality and Knowledge*.

Aubin, Jean-Pierre (1981). "Cooperative fuzzy games". In: *Mathematics of Operations Research* 6.1, pp. 1–13.

Aumann, Robert J and Michael Maschler (1964). "The bargaining set for cooperative games". In: *Advances in game theory* 52, pp. 443–476.

Aziz, Haris and Florian Brandl (2012). "Existence of stability in hedonic coalition formation games". In: *Proc of the 11th International Conference on Autonomous Agents and Multiagent Systems*, pp. 763–770.

Bahar, R Iris, Erica A Frohm, Charles M Gaona, Gary D Hachtel, Enrico Macii, Abelardo Pardo, and Fabio Somenzi (1997). "Algebraic decision diagrams and their applications". In: *Formal methods in system design* 10.2-3, pp. 171–206.

Balcan, Maria-Florina, Ariel D Procaccia, and Yair Zick (2015). "Learning cooperative games". In: *In Proc. of the 24th International Joint Conference on Artificial Intelligence*, pp. 475–481.

Barber, David (2012). *Bayesian reasoning and machine learning*. Cambridge University Press.

Bast, Holger and Ingmar Weber (2005). "Don't compare averages". In: *Experimental and Efficient Algorithms*. Springer, pp. 67–76.

Bird, Charles G (1976). "On cost allocation for a spanning tree: a game theoretic approach". In: *Networks* 6.4, pp. 335–350.

Bishop, C (2006). *Pattern Recognition and Machine Learning*. New York: Springer-Verlag.

Bistaffa, Filippo, Alessandro Farinelli, Jesús Cerquides, Juan Rodríguez-Aguilar, and Sarvapali D Ramchurn (2014). "Anytime coalition structure generation on synergy graphs". In: *Proc. of the 2014 international conference on Autonomous agents and multi-agent systems*, pp. 13–20.

Bistaffa, Filippo, Alessandro Farinelli, Georgios Chalkiadakis, and Sarvapali D Ramchurn (2015). "Recommending fair payments for large-scale social ridesharing". In: *Proceedings of the 9th ACM Conference on Recommender Systems*. ACM, pp. 139–146.

Blei, David M (2012). "Probabilistic topic models". In: *Communications of the ACM* 55.4, pp. 77–84.

Blei, David M, Andrew Y Ng, and Michael I Jordan (2003). "Latent dirichlet allocation". In: *Journal of Machine Learning Research* 3, pp. 993–1022.

Bodlaender, Hans L and Arie MCA Koster (2008). "Combinatorial optimization on graphs of bounded treewidth". In: *The Computer Journal* 51.3, pp. 255–269.

Bollobás, Béla (2001). "Random Graphs". In: *Cambridge Studies in Advanced Mathematics*.

Bondareva, Olga N (1963). "Some applications of linear programming methods to the theory of cooperative games". In: *Problemy kibernetiki* 10, pp. 119–139.

Boutilier, Craig, Yoav Shoham, and Michael P Wellman (1997). "Economic principles of multi-agent systems". In: *Artificial Intelligence* 94.1, pp. 1–6.

Brandt, Felix, Vincent Conitzer, Ulle Endriss, Jérôme Lang, and Ariel D Procaccia (2016). *Handbook of Computational Social Choice*. Cambridge University Press.

Bratman, Michael E, David J Israel, and Martha E Pollack (1988). "Plans and resource-bounded practical reasoning". In: *Computational intelligence* 4.3, pp. 349–355.

Caldara, Dario, Cristina Fuentes-Albero, Simon Gilchrist, and Egon Zakrajšek (2016). "The macroeconomic impact of financial and uncertainty shocks". In: *European Economic Review*.

Chalkiadakis, Georgios and Craig Boutilier (2004). "Bayesian reinforcement learning for coalition formation under uncertainty". In: *Proc. of the 3rd International Joint Conference on Autonomous Agents and Multiagent Systems*, pp. 1090–1097.

Chalkiadakis, Georgios and Craig Boutilier (2007). "Coalitional Bargaining with Agent Type Uncertainty". In: *IJCAI*, pp. 1227–1232.

Chalkiadakis, Georgios, Edith Elkind, and Michael Wooldridge (2011). *Computational aspects of cooperative game theory*. Morgan & Claypool Publishers.

Chalkiadakis, Georgios, Gianluigi Greco, and Evangelos Markakis (2016). "Characteristic function games with restricted agent interactions: Core-stability and coalition structures". In: *Artificial Intelligence* 232, pp. 76–113.

Chalkiadakis, Georgios, Evangelos Markakis, and Craig Boutilier (2007). "Coalition formation under uncertainty: Bargaining equilibria and the Bayesian core stability concept". In: *Proc. of the 6th International Joint Conference on Autonomous Agents and Multiagent Systems*, pp. 400–407.

Chalkiadakis, Georgios, Evangelos Markakis, and Nicholas R Jennings (2012). "Coalitional stability in structured environments". In: *Proc. of the 11th International Joint Conference on Autonomous Agents and Multiagent Systems*, pp. 779–786.

Chalkiadakis, Georgios, Edith Elkind, Evangelos Markakis, and Nicholas R Jennings (2008). "Overlapping Coalition Formation". In: *Proc. of the 4th International Workshop on Internet and Network Economics*, pp. 307–321.

Chalkiadakis, Georgios, Edith Elkind, Maria Polukarov, and Nicholas R Jennings (2009). "The price of democracy in coalition formation". In: *Proc. of The 8th International Conference on Autonomous Agents and Multiagent Systems*, pp. 401–408.

Chalkiadakis, Georgios, Edith Elkind, Evangelos Markakis, Maria Polukarov, and Nick R Jennings (2010). "Cooperative games with overlapping coalitions". In: *Journal of Artificial Intelligence Research* 39.1, pp. 179–216.

Chalkiadakis, Georgios, Valentin Robu, Ramachandra Kota, Alex Rogers, and Nicholas R Jennings (2011). "Cooperatives of distributed energy resources for efficient virtual power plants". In: *Proc. of the 10th International Conference on Autonomous Agents and Multiagent Systems*, pp. 787–794.

Chernoff, Herman (1952). "A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations". In: *The Annals of Mathematical Statistics*, pp. 493–507.

Claus, Caroline and Craig Boutilier (1998). "The dynamics of reinforcement learning in cooperative multiagent systems". In: *Proc. of the 15th National Conference on Artificial Intelligence (AAAI)*, pp. 746–752.

Cohen, Philip R and Hector J Levesque (1990). "Intention is choice with commitment". In: *Artificial intelligence* 42.2-3, pp. 213–261.

Conitzer, Vincent and Tuomas Sandholm (2003). "Complexity of determining nonemptiness of the core". In: *Proceedings of the 4th ACM Conference on Electronic Commerce*. ACM, pp. 230–231.

Conitzer, Vincent and Tuomas Sandholm (2006). "Complexity of constructing solutions in the core based on synergies among coalitions". In: *Artificial Intelligence* 170.6, pp. 607–619.

Cortes, Corinna and Vladimir Vapnik (1995). "Support-vector networks". In: *Machine learning* 20.3, pp. 273–297.

Dang, Viet Dung, Rajdeep K. Dash, Alex Rogers, and Nicholas R. Jennings (2006). "Overlapping coalition formation for efficient data fusion in multi-sensor networks". In: *Proc. of the 21st National Conference on Artificial Intelligence (AAAI)*, pp. 635–640.

Daskalakis, Constantinos and Christos H Papadimitriou (2015). "Approximate Nash equilibria in anonymous games". In: *Journal of Economic Theory* 156, pp. 207–245.

Davis, Morton and Michael Maschler (1965). "The kernel of a cooperative game". In: *Naval Research Logistics Quarterly* 12.3, pp. 223–259.

Dechesne, Francien, Gennaro Di Tosto, Virginia Dignum, and Frank Dignum (2013). "No smoking here: values, norms and culture in multi-agent systems". In: *Artificial intelligence and law* 21.1, pp. 79–107.

Deng, Xiaotie, Toshihide Ibaraki, and Hiroshi Nagamochi (1999). "Algorithmic aspects of the core of combinatorial optimization games". In: *Mathematics of Operations Research* 24.3, pp. 751–766.

Deng, Xiaotie and Christos H Papadimitriou (1994). "On the complexity of cooperative solution concepts". In: *Mathematics of Operations Research* 19.2, pp. 257–266.

Eberhart, Russ C, James Kennedy, et al. (1995). "A new optimizer using particle swarm theory". In: *Proc. of the 6th international symposium on micro machine and human science*. Vol. 1. New York, NY, pp. 39–43.

Elkind, Edith, Dmitrii Pasechnik, and Yair Zick (2013). "Dynamic weighted voting games". In: *Proc. of the 12th International Conference on Autonomous Agents and Multi-Agent Systems*, pp. 515–522.

Elkind, Edith and Michael Wooldridge (2009). "Hedonic coalition nets". In: *Proceedings of The 8th International Conference on Autonomous Agents and Multiagent Systems-Volume 1*. International Foundation for Autonomous Agents and Multiagent Systems, pp. 417–424.

Franco, Pedro (2014). *Understanding Bitcoin: Cryptography, engineering and economics*. John Wiley & Sons.

Fudenberg, Drew and David K Levine (1998). *The theory of learning in games*. MIT press.

Garey, Michael R. and David S. Johnson (1979). *Computers and Intractability: A Guide to the Theory of NP-Completeness*. New York, NY, USA: W. H. Freeman & Co. ISBN: 0716710447.

Gillies, Donald B (1959). "Solutions to general non-zero-sum games". In: *Contributions to the Theory of Games* 4.40, pp. 47–85.

Glynn, Peter W and Dirk Ormoneit (2002). "Hoeffding's inequality for uniformly ergodic Markov chains". In: *Statistics & probability letters* 56.2, pp. 143–146.

Goemans, Michel X and Martin Skutella (2004). "Cooperative facility location games". In: *Journal of Algorithms* 50.2, pp. 194–214.

Grimmett, Geoffrey and David Stirzaker (2001). *Probability and random processes*. Oxford University Press.

Hoeffding, Wassily (1963). "Probability inequalities for sums of bounded random variables". In: *Journal of the American Statistical Association* 58, pp. 13–30.

Hoffman, Matthew, Francis R Bach, and David M Blei (2010). "Online learning for latent dirichlet allocation". In: *Advances in Neural Information Processing Systems*, pp. 856–864.

Hofmann, Thomas (1999). "Probabilistic latent semantic indexing". In: *Proceedings of the 22nd annual international ACM SIGIR conference on Research and development in information retrieval*, pp. 50–57.

Hu, Diane and Lawrence K Saul (2009). "A Probabilistic Topic Model for Unsupervised Learning of Musical Key-Profiles." In: *ISMIR*, pp. 441–446.

Hu, Junling and Michael P Wellman (1998). "Multiagent reinforcement learning: theoretical framework and an algorithm." In: *Proceedings of the 15th International Conference on Machine Learning*, pp. 242–250.

Hu, Junling and Michael P Wellman (2003). "Nash Q-learning for general-sum stochastic games". In: *Journal of Machine Learning Research*, pp. 1039–1069.

Ieong, Samuel and Yoav Shoham (2005). "Marginal contribution nets: a compact representation scheme for coalitional games". In: *Proceedings of the 6th ACM conference on Electronic commerce*, pp. 193–202.

Ieong, Samuel and Yoav Shoham (2008). "Bayesian Coalitional Games." In: *Proc. of the 23rd National Conference on Artificial Intelligence (AAAI)*, pp. 95–100.

Igarashi, Ayumi and Edith Elkind (2016). "Hedonic games with graph-restricted communication". In: *Proceedings of the 2016 International Conference on Autonomous Agents & Multiagent Systems*, pp. 242–250.

Impagliazzo, Russell (1995). "A personal view of average-case complexity". In: *Structure in Complexity Theory Conference, 1995, Proceedings of 10th Annual IEEE*, pp. 134–147.

Jolliffe, Ian (2002). *Principal component analysis*. Wiley Online Library.

Jordan, Michael I, Zoubin Ghahramani, Tommi S Jaakkola, and Lawrence K Saul (1999). "An introduction to variational methods for graphical models". In: *Machine learning* 37.2, pp. 183–233.

Jordan, Michael Irwin (1998). *Learning in graphical models*. Springer Science & Business Media.

Kaisers, Michael and Karl Tuyls (2009). "Replicator dynamics for multi-agent learning: an orthogonal approach". In: *International Workshop on Adaptive and Learning Agents*. Springer, pp. 49–59.

Kalai, Ehud and Eitan Zemel (1982a). "Generalized network problems yielding totally balanced games". In: *Operations Research* 30.5, pp. 998–1008.

Kalai, Ehud and Eitan Zemel (1982b). "Totally balanced games and games of flow". In: *Mathematics of Operations Research* 7.3, pp. 476–478.

Kapetanakis, Spiros and Daniel Kudenko (2002). "Reinforcement learning of coordination in cooperative multi-agent systems". In: *Proceedings of the 18th National Conference on Artificial Intelligence (AAAI)*, pp. 326–331.

Kellerer, Hans, Ulrich Pferschy, and David Pisinger (2004). *Knapsack problems*. Springer, Berlin.

Koller, Daphne and Nir Friedman (2009). *Probabilistic graphical models: principles and techniques*. MIT press.

Komatsu, Takao (2003). "On the number of solutions of the Diophantine equation of Frobenius-General case". In: *Mathematical Communications* 8.2, pp. 195–206.

Koutsoupias, Elias and Christos Papadimitriou (1999). "Worst-case equilibria". In: *Annual Symposium on Theoretical Aspects of Computer Science*. Springer, pp. 404–413.

Kraus, Sarit, Onn Shehory, and Gilad Taase (2003). "Coalition formation with uncertain heterogeneous information". In: *Proc. of the 2nd International Joint Conference on Autonomous Agents and Multiagent Systems*, pp. 1–8.

Kraus, Sarit, Onn Shehory, and Gilad Taase (2004). "The advantages of compromising in coalition formation with incomplete information". In: *Proc. of the 3rd International Joint Conference on Autonomous Agents and Multiagent Systems*, pp. 588–595.

Leskovec, Jure and Andrej Krevl (2014). *SNAP Datasets: Stanford Large Network Dataset Collection*. http://snap.stanford.edu/data.

Li, Cuihong, Shuchi Chawla, Uday Rajan, and Katia Sycara (2005). "Mechanism design for coalition formation and cost sharing in group-buying markets". In: *Electronic Commerce Research and Applications* 3.4, pp. 341–354.

Littman, Michael L (1994). "Markov games as a framework for multi-agent reinforcement learning". In: *Proceedings of the 11th International Conference on Machine Learning*, pp. 157–163.

Liu, Bing, Lin Liu, Anna Tsykin, Gregory J Goodall, Jeffrey E Green, Min Zhu, Chang Hee Kim, and Jiuyong Li (2010). "Identifying functional miRNA–mRNA regulatory modules with correspondence latent dirichlet allocation". In: *Bioinformatics* 26.24, pp. 3105–3111.

Maskeri, Girish, Santonu Sarkar, and Kenneth Heafield (2008). "Mining business topics in source code using latent dirichlet allocation". In: *Proceedings of the 1st India software engineering conference*. ACM, pp. 113–120.

Massart, Pascal (2000). "Some applications of concentration inequalities to statistics". In: *Annales de la Faculté des sciences de Toulouse: Mathématiques*. Vol. 9. 2, pp. 245–303.

Michalski, Ryszard S, Jaime G Carbonell, and Tom M Mitchell (2013). *Machine learning: An artificial intelligence approach*. Springer Science & Business Media.

Mitzenmacher, Michael and Eli Upfal (2005). *Probability and computing: Randomized algorithms and probabilistic analysis*. Cambridge University Press.

Motwani, Rajeev and Prabhakar Raghavan (1995). *Randomized algorithms*. New York: Cambridge Univ. Press.

Myerson, Roger B (1977). "Graphs and cooperation in games". In: *Mathematics of Operations Research* 2.3, pp. 225–229.

Myerson, Roger B (2013). *Game theory*. Harvard university press.

Nash, John (1951). "Non-cooperative games". In: *Annals of mathematics*, pp. 286–295.

Nigam, Kamal, Andrew Kachites McCallum, Sebastian Thrun, and Tom Mitchell (2000). "Text classification from labeled and unlabeled documents using EM". In: *Machine learning* 39.2-3, pp. 103–134.

Nisan, Noam, Tim Roughgarden, Eva Tardos, and Vijay V Vazirani (2007). *Algorithmic game theory*. Vol. 1. Cambridge University Press Cambridge.

Ohta, Naoki, Atsushi Iwasaki, Makoto Yokoo, Kohki Maruono, Vincent Conitzer, and Tuomas Sandholm (2006). "A compact representation scheme for coalitional games in open anonymous environments". In: *AAAI*. Vol. 6, 21st.

Oliveira Ramos, Gabriel de and Ana LC Bazzan (2012). "Reduction of Coalition Structure's Search Space Based on Domain Information: An Application in Smart Grids". In: *Social Simulation (BWSS), 2012 Third Brazilian Workshop on*. IEEE, pp. 112–119.

Osborne, Martin J and Ariel Rubinstein (1994). *A course in game theory*. MIT press.

Paley, Raymond and Antoni Zygmund (1932). "A note on analytic functions in the unit circle". In: *Mathematical Proceedings of the Cambridge Philosophical Society*. Vol. 28. 03. Cambridge Univ Press, pp. 266–272.

Papadimitriou, Christos (2001). "Algorithms, games, and the internet". In: *Proc. of the 33rd Annual ACM Symposium on Theory of Computing*, pp. 749–753.

Proakis, John G and Dimitris G Manolakis (1996). *Digital Signal Processing: Principles, Algotithms, and Applications*. Prentice Hall.

Rahwan, Talal and Nicholas R Jennings (2008). "An improved dynamic programming algorithm for coalition structure generation". In: *Proceedings of the 7th international joint conference on Autonomous agents and multiagent systems*, pp. 1417–1420.

Rahwan, Talal, Sarvapali D Ramchurn, Nicholas R Jennings, and Andrea Giovannucci (2009). "An anytime algorithm for optimal coalition structure generation". In: *Journal of Artificial Intelligence Research* 34, pp. 521–567.

Rahwan, Talal, Tomasz P Michalak, Michael Wooldridge, and Nicholas R Jennings (2015). "Coalition structure generation: A survey". In: *Artificial Intelligence* 229, pp. 139–174.

Rao, Anand S and Michael P Georgeff (1991). "Modeling rational agents within a BDI-architecture." In: *KR* 91, pp. 473–484.

Rapoport, Anatol (1970). *N-person game theory: Concepts and applications*. Courier Corporation.

Robu, Valentin, Ramachandra Kota, Georgios Chalkiadakis, Alex Rogers, and Nicholas R Jennings (2012). "Cooperative virtual power plant formation using scoring rules". In: *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems*, pp. 1165–1166.

Russell, Stuart and Peter Norvig (1995). *Artificial Intelligence: A Modern Approach*. Prentice Hall.

Sandholm, Tuomas, Kate Larson, Martin Andersson, Onn Shehory, and Fernando Tohmé (1999). "Coalition structure generation with worst case guarantees". In: *Artificial Intelligence* 111.1, pp. 209–238.

Schmeidler, David (1969). "The nucleolus of a characteristic function game". In: *SIAM Journal on applied mathematics* 17.6, pp. 1163–1170.

Sebastiani, Fabrizio (2002). "Machine learning in automated text categorization". In: *ACM computing surveys* 34.1, pp. 1–47.

Shapley, Lloyd S (1967). "On balanced sets and cores". In: *Naval Research Logistics (NRL)* 14.4, pp. 453–460.

Shapley, LS (1952). "A value for n-person games". In: *Contributions to the Theory of Games* 2, p. 307.

Shehory, Onn and Sarit Kraus (1998). "Methods for task allocation via agent coalition formation". In: *Artificial Intelligence* 101.1, pp. 165–200.

Shehory, Onn and Sarit Kraus (1999). "Feasible formation of coalitions among autonomous agents in nonsuperadditive environments". In: *Computational Intelligence* 15.3, pp. 218–251.

Shoham, Yoav and Kevin Leyton-Brown (2008). *Multiagent systems: Algorithmic, game-theoretic, and logical foundations*. Cambridge University Press.

Shoham, Yoav, Rob Powers, and Trond Grenager (2007). "If multi-agent learning is the answer, what is the question?" In: *Artificial Intelligence* 171.7, pp. 365–377.

Stone, Peter (2007). "Multiagent learning is not the answer. It is the question". In: *Artificial Intelligence* 171.7, pp. 402–405.

Suijs, Jeroen, Peter Borm, Anja De Waegenaere, and Stef Tijs (1999). "Cooperative games with stochastic payoffs". In: *European Journal of Operational Research* 113.1, pp. 193–205.

Sutton, Richard S and Andrew G Barto (1998). *Reinforcement learning: An introduction*. MIT press Cambridge.

Teh, Yee Whye, Michael I Jordan, Matthew J Beal, and David M Blei (2006). "Hierarchical dirichlet processes". In: *Journal of the American Statistical Association*, pp. 1566–1581.

Tuyls, Karl and Simon Parsons (2007). "What evolutionary game theory tells us about multiagent learning". In: *Artificial Intelligence* 171.7, pp. 406–416.

Tuyls, Karl and Gerhard Weiss (2012). "Multiagent learning: Basics, challenges, and prospects". In: *AI Magazine* 33.3, p. 41.

Ueda, Suguru, Makoto Kitaki, Atsushi Iwasaki, and Makoto Yokoo (2011). "Concise characteristic function representations in coalitional games based on agent types". In: *Proc. of the 22nd International Joint Conference on Artificial Intelligence*, pp. 393–399.

Vazirani, Vijay (2001). *Approximation Algorithms*. Springer.

Voice, Thomas, Maria Polukarov, and Nicholas R Jennings (2012). "Coalition structure generation over graphs". In: *Journal of artificial intelligence research*.

Voice, Thomas, Sarvapali D Ramchurn, and Nicholas R Jennings (2012). "On coalition formation with sparse synergies". In: *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems*, pp. 223–230.

Von Neumann, John and Oskar Morgenstern (1947). *Theory of games and economic behavior*. Princeton university press.

Wang, Tianyu, Lingyang Song, Zhu Han, and Walid Saad (2014). "Distributed cooperative sensing in cognitive radio networks: An overlapping coalition formation approach". In: *IEEE Transactions on Communications* 62.9, pp. 3144–3160.

Watkins, Christopher JCH and Peter Dayan (1992). "Q-learning". In: *Machine learning* 8, pp. 279–292.

Weerdt, Mathijs de, Yingqian Zhang, and Tomas Klos (2007). "Distributed task allocation in social networks". In: *Proc. of the 6th International Conference on Autonomous Agents and Multi-Agent Systems*, pp. 500–507.

Wooldridge, Michael (2009). *An introduction to multiagent systems*. John Wiley & Sons.

Wooldridge, Michael and Nicholas R Jennings (1995). "Intelligent agents: Theory and practice". In: *The knowledge engineering review* 10.02, pp. 115–152.

Yamamoto, Junichi and Katia Sycara (2001). "A stable and efficient buyer coalition formation scheme for e-marketplaces". In: *Proceedings of the 5th International Conference on Autonomous agents*. ACM, pp. 576–583.

Zhang, Guofu, Jianguo Jiang, Zhaopin Su, Meibin Qi, and Hua Fang (2010). "Searching for overlapping coalitions in multiple virtual organizations". In: *Information Sciences* 180.17, pp. 3140–3156.

Zick, Yair, Georgios Chalkiadakis, and Edith Elkind (2012). "Overlapping coalition formation games: Charting the tractability frontier". In: *Proc. of the 11th International Joint Conference on Autonomous Agents and Multiagent Systems*, pp. 787–794.

Zick, Yair and Edith Elkind (2011). "Arbitrators in overlapping coalition formation games". In: *The 10th International Conference on Autonomous Agents and Multiagent Systems-Volume*, pp. 55–62.

Zick, Yair, Evangelos Markakis, and Edith Elkind (2014). "Arbitration and stability in cooperative games with overlapping coalitions". In: *Journal of Artificial Intelligence Research* 50, pp. 847–884.