BEN-GURION UNIVERSITY OF THE NEGEV

# Communication Awareness in Multi-Agent Optimization

*Author:*
Ben Rachmut

*Supervisor:*
Prof. Roie Zivan

Thesis submitted in partial fulfillment of the requirements for the degree of "Doctor of Philosophy"

Submitted to the Senate of Ben-Gurion University of the Negev
March 31, 2024

Beer-Sheva

אוניברסיטת בן גוריון בנגב

# מודעות תקשורת באופטימיזציה של מערכות מרובות סוכנים


מאת:
בן רחמוט

בהנחיית:
פרופ' רועי זיוון


מחקר לשם מילוי חלקי של הדרישות לקבלת תואר

"דוקטור לפילוסופיה"

הוגש לסינאט אוניברסיטת בן גוריון בנגב

כ"א באדר ב
תשפד

31 במרץ
2024

באר שבע

Ben-Gurion University of the Negev

# Communication Awareness in Multi-Agent Optimization

*Author:*
Ben Rachmut

*Supervisor:*
Prof. Roie Zivan

Thesis submitted in partial fulfillment of the requirements for the degree of "Doctor of Philosophy"

Submitted to the Senate of Ben-Gurion University of the Negev

Approved by the advisor: ____________

Approved by the Dean of the Kreitman School of Advanced Graduate Studies: ____________

March 31, 2024

Beer-Sheva

# Declaration of Authorship

This work was carried out under the supervision of
Prof. Roie Zivan
In the Department of Industrial Engineering and Management
Faculty of Engineering Sciences

I, Ben Rachmut, whose signature appears below, hereby declare that this thesis titled, "Communication Awareness in Multi-Agent Optimization" and the work presented in it are my own. I confirm that:

- I have written this Thesis by myself, except for the help and guidance offered by my Thesis Advisors.
- The scientific materials included in this Thesis are products of my own research, culled from the period during which I was a research student.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given.

Student's name: Ben Rachmut

---

Signature:

---

Date:

---

# *Acknowledgements*

I would like to acknowledge everyone who has assisted me throughout my doctoral studies over the years.

# *Abstract*


Most studies investigating Multi-Agent Optimization (MAO) assume that messages arrive instantaneously and are never lost. Distributed Constraint Optimization Problems (DCOPs) and General Task Allocation Problems (GTAPs) are examples of models used to solve real-life MAO problems; however, they do not consider the effects of imperfect communication. In distributed systems, in which communication is a key factor, agents rely on message delivery to achieve mutual goals. However, the assumption of seamless communication is not always practical because it can be subject to delays and failures. To address this gap in the existing models, we introduce Communication Aware Multi-agent Optimization (CA-MAO) models. These models were specifically designed to incorporate dynamic uncertainties in communication within distributed systems. By extending the models to their communication-aware versions, we enhance their ability to represent real-world distributed settings where communication challenges are accounted for.

Many state-of-the-art MAO algorithms have been designed synchronously, despite running in asynchronous environments. These algorithms rely on synchronous iterations, wherein each agent interacts with every neighbor in each cycle, the challenge arises from the uncertainty of message delivery. The reliance on message delivery creates a significant challenge for such synchronous algorithmic design. Uncertainty in message delivery can cause delays or even deadlocks, as agents may become stuck waiting for messages that will never arrive. These issues can significantly impact the efficiency and effectiveness of the algorithm.

In addition, some of these algorithms and mechanisms take advantage of common simplistic communication assumptions to achieve desirable properties, such as monotonicity and convergence. To address the limitations of current algorithms that fail to consider the effects of imperfect communication, we propose communication-aware algorithms.

This study advances the field of MAO research in four ways:

1. First we explore distributed local search algorithms for CA-DCOPs. We evaluated the performance of the existing synchronous algorithms and developed their asynchronous counterparts. We also propose a novel asynchronous anytime framework to enhance exploration in nonmonotonic asynchronous algorithms. Our results demonstrate that imperfect communication can surprisingly improve distributed local search algorithms, leading to increased exploration.

2. Our second contribution is focused on distributed local search algorithms that guarantee solutions that cannot be improved by a group of k agents, known as k-opt solutions.

We present the Latency-Aware Monotonic Distributed Local Search (LAMDLS) and LAMDLS-2 algorithms: latency aware search algorithms for solving DCOPs. These algorithms are monotonic and guarantee 1-opt and 2-opt convergence, respectively. Additionally, they exhibit resilience to message latency.

3. In our third contribution, we explored inference algorithms for CA-DCOPs, with a particular focus on the Max-sum algorithm and its variants. We aimed to assess the performance of synchronous and asynchronous versions of the algorithm in various communication scenarios through both empirical and theoretical analyses. Our findings indicate that the damped Max-sum algorithm exhibits remarkable robustness when faced with imperfect communication.

4. The fourth contribution introduces the CA-GTAPs. We focus on the Fisher Market Clearing Task Allocation (FMC_TA) algorithm, which outperforms centralized and distributed optimization algorithms but is not suitable for imperfect communication scenarios. We propose an asynchronous FMC_ATA algorithm that demonstrates resistance to imperfect communication without compromising solution quality. Our investigation also shows that the distributed version of the algorithm is preferred over the centralized version when communication is extremely poor and produces consistent results.

**Keywords: Multi-Agent Optimization, Communication-Awareness, Distributed Constraint Optimization, Task Allocation, Asynchronous Algorithms, Imperfect Communication**

# תקציר

רוב המחקרים בתחום אופטימיזציה מרובת סוכנים (MAO) מניחים שהודעות מגיעות מיידית ולא הולכות לאיבוד. בעיות אופטימיזציה מבוזרות של אילוצים (DCOPs) ובעיות הקצאת משימות כלליות (GTAPs) הן דוגמאות למודלים המשמשים לפתרון בעיות MAO מציאותיות. עם זאת, מודלים אלו לא מתחשבים בהשפעות של הפרעות בתקשורת. במערכות מבוזרות, בהן תקשורת היא גורם מפתח, סוכנים מסתמכים על העברת הודעות כדי להשיג מטרות משותפות. עם זאת, ההנחה של תקשורת מושלמת לא תמיד מציאותית, שכן היא עלולה להיות נתונה לעיכובים וכשלים. כדי להתמודד עם פער זה במודלים הקיימים, אנו מציגים מודלים של אופטימיזציה מרובת סוכנים עם מודעות תקשורת מודלים אלה תוכננו במיוחד כדי לשלב חוסר ודאות דינמית בתקשורת, בתוך מערכות מבוזרות. הרחבת המודלים מאפשרת שיפור של יכולתם לייצג בעיות מבוזרות בעולם האמיתי המתחשבות באתגרי תקשורת.

רבים מהאלגוריתמים עבור פתרון בעיותMAO תוכננו באופן סינכרוני, למרות שהם פועלים בסביבה אסינכרונית. האלגוריתמים הללו מסתמכים על איטרציות סינכרוניות, בהן סוכן מתקשר עם כל שכן בכל מחזור. הסתמכות על העברת הודעות יוצרת אתגר משמעותי עבור עיצוב סינכרוני. אי ודאות בהעברת הודעות עלולה לגרום לעיכובים ואף למבוי סתום, שכן סוכנים עלולים להיתקע בהמתנה להודעות שלא יגיעו לעולם. בעיות אלו יכולות להשפיע באופן משמעותי על היעילות והאפקטיביות של האלגוריתם.

מחקר זה מקדם את תחום חקר ה-MAO בארבע דרכים:

1. חקר אלגוריתמי חיפוש מקומיים מבוזרים עבור בעיות אופטימיזציה מבוזרות עם אילוצים: הערכנו את הביצועים של האלגוריתמים הסינכרוניים הקיימים ופיתחנו את גרסאותיהם האסינכרוניות. בנוסף, אנו מציעים מסגרת חדשה, אסינכרונית, לקבלת תכונת ה,anytime על מנת לאפשר חקירה במרחב הפתרונות (exploration) עבור אלגוריתמים אסינכרוניים לא מונוטוניים. התוצאות שלנו מראות שתקשורת לא מושלמת יכולה לשפר באופן מפתיע אלגוריתמי חיפוש מקומיים מבוזרים באמצעות הגברת החקירה של מרחב הפתרונות.
2. התרומה השנייה שלנו מתמקדת באלגוריתמי חיפוש מקומיים מבוזרים, המבטיחים פתרונות שאינם ניתנים לשיפור על ידי קבוצה של K סוכנים, הידועים כפתרונות K-OPT. אנו מציגים את ה (Latency-Aware Monotonic Distributed Local Search (LAMDLS ואת 2-LAMDLS אלגוריתמי חיפוש עבור בעיות אופטימיזציה מבוזרות עם אילוצים ותקשורת מודעת. אלגוריתמים אלו מבטיחים התכנסות של opt-1 ו opt-2 בהתאמה.

3. התרומה השלישית שלנו מתמקדת בחקר אלגוריתמי הסקה עם דגש על אלגוריתם ה Max sum והגרסאות השונות שלו, עבור בעיות אופטימיזציה מבוזרות עם אילוצים בסביבה עם בעיות תקשורת. באמצעות ניתוחים אמפיריים ותיאורטיים, בחנו את הביצועים של גרסאות סינכרוניות ואסינכרוניות של האלגוריתם בתרחישי תקשורת שונים. הממצאים שלנו מצביעים על כך שאלגוריתם הdamp Max sum מפגין עמידות יוצאת דופן.
4. לבסוף הצגנו את בעיות הקצאת משימות הכללית עם בעיות תקשורת. אנו התמקדנו באלגוריתם Fisher Market Clearing Task Allocation המפגין ביצועים טובים יותר מאלגוריתמי אופטימיזציה ריכוזיים ומבוזרים רבים אך אינו מתאים לתרחישי תקשורת לא מושלמת. לצורך התמודדות עם אתגר זה, אנו מציעים אלגוריתם אסינכרוני בשם FMC_ATA המפגין חסינות לבעיות תקשורת ללא התפשרות על איכות הפתרון. המחקר שלנו גם מראה שהגרסה המבוזרת של האלגוריתם עדיפה על הגרסה הריכוזית, כאשר התקשורת גרועה מאוד, והיא מפיקה תוצאות עקביות.

מילות מפתח: אופטימיזציה מרובת סוכנים, מודעות לתקשורת, אופטימיזציה של אילוצים מבוזרים, הקצאת משימות, אלגוריתמים אסינכרוניים, תקשורת לא מושלמת.

# Contents

# List of Figures

# List of Tables

# List of Abbreviations

| | |
|---|---|
| **CA** | **C**ommunication **A**ware |
| **MAO** | **M**ulti-**A**gent **O**ptimization |
| **CCG** | **C**onstrained **C**ommunication **G**raph |
| **DALO** | **D**istributed **A**synchronous **L**ocal **O**ptimization |
| **DCOP** | **D**istributed **C**onstraint **O**ptimization **P**roblem |
| **DSA** | **D**istributed **S**tochastic **A**lgorithm |
| **DSA-SDP** | **D**istributed **S**tochastic **A**lgorithm- **S**lope **D**ependent **P**robability |
| **FMC_ATA** | **F**isher **M**arket **C**learing **A**synchronous **T**ask **A**llocation |
| **FMC_TA** | **F**isher **M**arket **C**learing **T**ask **A**llocation |
| **GTAP** | **G**eneral **T**ask **A**llocation **P**roblem |
| **LAMDLS** | **L**atency **A**ware **M**onotonic **D**istributed **L**ocal **S**earch |
| **MAS** | **M**ulti-**A**gent **S**ystems |
| **MGM** | **M**aximum **G**ain **M**essages |

# Chapter 1

# Introduction

Recent advancements in computation and communication have resulted in realistic distributed applications in which humans and technology interact and aim to optimize mutual goals(e.g., IoT applications). Thus, there is a growing need for optimization methods to support decentralized decision-making in complex Multi-Agent Systems (MAS), which find applications in various domains, including device scheduling in smart homes, target tracking in sensor networks, mission planning for unmanned autonomous vehicles, and coordination of rescue units in disaster scenarios, as noted in [10, 22, 32].

To comprehend the complexities and challenges associated with MAS, we considered the disaster management domain. In such scenarios, diverse rescue units such as medical staff, firefighters, and law enforcement must coordinate their actions to save as many victims as possible. Such a multi-agent task coordination problem is challenging due to the following characteristics.

- **Optimization of a Global Objective:** The various rescue units need to work together towards a common goal (e.g., maximizing the number of victims saved). One such instance is when police units are required to facilitate the safe passage of ambulances carrying victims out of a disaster area.

- **Decentralized Coordination:** There is often not a centralized entity that manages the coordination problem. For example, a hospital administrator will want to manage its medical personnel, a fire chief will want to manage its firefighters, and a police chief will want to manage its police units. If there is a larger incident that requires units from multiple hospitals, fire stations, and police stations, then more decision-makers will need to coordinate with each other to identify the best way to deploy the combined rescue units.

- **Imperfect Communication:** Finally, the quality of communication in such scenarios can significantly deteriorate. For instance, a disaster may damage the cell transmission towers that various rescue units rely on for communication. Consequently, standard algorithms for coordination problems that rely on perfect communication may be impractical.

Although we use the disaster rescue scenario as the motivating domain, each of these factors is also present in a larger class of other Multi-Agent Optimization (MAO) problems.

Well-known multi-agent approaches, such as Distributed Constraint Optimization Problems (DCOPs)[8, 16, 23] and the General Task Allocation Problem (GTAP)[24, 25] effectively address key aspects of decentralized decision-making. Both models utilize agents to represent decision-makers, i.e., rescue units. In DCOPs, agents assign values to their variables, as the rescue unit chief assigns duties to the personnel. GTAP is similar to DCOPs in that it assigns agents to tasks. However, GTAP introduces additional challenges imposed by realistic applications, including spatial and temporal constraints. Consequently, in this context, the rescue unit chief must create schedules for each personnel and establish ad hoc coalitions to ensure that the tasks are completed efficiently.

Both the multi-agent optimization problems mentioned above aim to optimize a global objective through decentralized coordination, thereby capturing the first two characteristics (i.e., global optimization and decentralized coordination). To achieve this coordination, agents communicate and coordinate their actions through a message exchange. Unfortunately, the communication assumptions of these models are overly simplistic and often unrealistic: (1) messages are never lost, (2) all messages have very small and bounded delays, and (3) the messages sent arrive in the order that they were sent. These assumptions do not reflect real-world characteristics, where messages may be disproportionally delayed or dropped because of congestion or different bandwidths in different communication channels.

Because such multi-agent optimization problems are NP-hard [16, 17], considerable research effort has been devoted to developing algorithms to find good solutions quickly. This class of algorithms, known as incomplete algorithms, has been the focus of numerous studies [1, 5, 11, 15, 18, 34, 39, 40, 6, 12, 17, 30, 33, 36]. Despite offering little or no quality guarantees, these algorithms have been empirically found to produce high-quality solutions. Examples of such algorithms include the Distributed Stochastic Algorithm (DSA), Maximum Gain Messages (MGM), and Max-Sum [15, 39, 40, 2, 3] for solving DCOPs and Fisher Market Clearing Task Allocation (FMC_TA) [17] for solving GTAPs.

The presence of imperfect communication can have a significant impact on the quality guarantees and desirable properties of such algorithms. Some of these algorithms do guarantee other desirable properties, for example, MGM-k guarantees monotonicity and convergence to a k-opt solution (i.e, a solution that cannot be improved by an action of a single agent [15, 20]). Another property that can be guaranteed is the anytime property. This property can be achieved using the anytime framework proposed in Zivan, Okamoto, and Peled [41]. Unfortunately, algorithms and the anytime framework take advantage of the common simplistic communication assumptions discussed above. Consequently, the guarantees for achieving these properties may no longer hold when communication is unreliable.

The general design of most state-of-the-art MAO algorithms is synchronous. However, the settings in which agents are expected to perform are asynchronous [14] because the environment in which they perform is distributed, and the agents do not hold a mutual clock. Therefore, synchronization was achieved by exchanging the messages in each iteration of the algorithm. In each iteration, an agent receives messages from its neighbors in the previous iteration, computes, and sends messages to all of its neighbors [40, 41]. Unfortunately, such a synchronous algorithmic design has several drawbacks when imperfect communication is considered. In the presence of message latency, every synchronous iteration is completed only after all messages sent in the previous iteration arrive. Therefore, the advancement of the algorithm from one iteration to the next depends on the longest message delay in that iteration. When messages can be lost, an agent may expect to receive a message from its neighbor, while the neighbor is unaware that the message it sends does not arrive. Thus, these agents are deadlocked, with each waiting for a message from the other.

This thesis seeks to address the challenges posed by Communication Aware Multi-Agent Optimization (CA-MAO) problems by developing incomplete algorithms that can adapt to varying communication reliabilities. Specifically, we focus on two approaches that represent such problems: DCOPs, which provide an abstract representation of multi-agent coordination, and GTAPs, which address a broader range of challenges typically encountered in real-life applications. The main tasks of this thesis are 1) Model Design, which involves creating communication-aware multi-agent optimization models that can represent dynamic communication uncertainties; thus, we introduce CA-DCOP and CA-GTAP; 2) Algorithm Design, which includes developing communication-aware algorithms that can address communication characteristics; and 3) Analytical and Empirical Evaluations, which assess the efficiency of the algorithms in terms of computational requirements, communication costs, solution quality, and convergence.

Chapter 2 provides the foundation for the DCOP and GTAP models, detailing their background and relevant information. In addition, the chapter outlines the benchmark algorithms and methods used for problem-solving, as well as related work on imperfect communication in multi-agent systems.

In Chapter 3, we address the limitations of DCOPs when communication is imperfect by presenting the Communication-Aware DCOP model (CA-DCOP), investigate the consequences of message latency and loss on synchronous benchmark local search algorithms, and propose an asynchronous approach for these algorithms. In addition, we developed an asynchronous anytime framework that allows for the best solution explored in non-monotonic asynchronous local search DCOP algorithms.

Chapter 4 examines the susceptibility of message latency to properties ensured by the MGM-k algorithm. We introduce Latency-Aware Monotonic Distributed Local Search (LAMDLS) and LAMDLS-2, which are algorithms that guarantee monotonicity and 1-opt and 2-opt solutions, respectively, upon convergence.

In 5, the focus was on examining inference algorithms in the context of CA-DCOPs. This study provided a comprehensive analysis of both the synchronous and alternative asynchronous designs, taking into account the presence of imperfect communication in the max-sum algorithm and its variants.

Chapter 6 provides an evaluation of the influence of message latency and loss on a real-world domain (i.e., GTAP) by introducing the Communication-Aware GTAP model (CA-GTAP). We focused on the FMC_TA algorithm and addressed its limitations. We propose FMC_ATA, an asynchronous version of FMC_TA, which is robust to message latency and loss and is more applicable in such scenarios. Finally, we investigated the conditions under which the distributed version of the algorithm was preferred to the centralized version.

Chapter 7 summarizes this dissertation, indicating its contributions and future work.

# Chapter 2

# Background

In the background chapter, we provide details of existing Multi-Agent Optimization (MAO) models and algorithms that presume perfect communication. This chapter delves into the following models: DCOPs and GTAPs. In addition, we review prior research on communication awareness in Multi-Agent Systems and the methodologies used to evaluate the performance of algorithms in scenarios with imperfect communication.

Within the realm of DCOP algorithms, we delve into their classification, with a particular focus on local search and inference as solution strategies [9]. In the context of local search algorithms, we delve into the Distributed Stochastic Algorithm (DSA) [40] and its variant, DSA with Slope-Dependent Probability (DSA-SDP), which moderately improves DSA's exploratory capabilities [41]. Additionally, we address an existing anytime mechanism that can be utilized in conjunction with distributed synchronous local search algorithms to maintain a record of the best solutions [41].

Subsequently, we delve into k-opt and Region-opt Algorithms [19, 35]. Specifically, we focus on Maximum Gain Messages-k (MGM-K) [15] and Distributed Asynchronous Local Optimization-k (DALO-k). These algorithms guarantee that the solutions obtained are $k$-opt [13], implying that they cannot be improved by a group of $k$ agents. MGM-K is designed as a synchronous algorithm, whereas DALO-K operates asynchronously.

For inference algorithms, we focused on adaptations of belief propagation [21, 38] for DCOPs [2, 3]. These include the Max-sum algorithm and its variants, such as Damped Max-sum and Max-sum with Split Constraint Factor Graph [3]. In addition, we provide details of the backtrack cost tree (BCT) to evaluate the performance of these algorithms [42].

For the GTAP model (an abstraction of the Law Enforcement Problem (LEP), which serves as a specific instance of GTAP as detailed in [17]), we shifted our focus to methodologies aimed at task allocation and optimization within complex real-world scenarios. We provide details of both the static and dynamic versions of the model and introduce the FMC_TA algorithm, as proposed in [17], highlighting its effectiveness in addressing such problems.

# Chapter 3

# Communication-Aware Local Search for Distributed Constraint Optimization

The impact of the synchronous design on the performance of distributed local-search algorithms for DCOPs in the face of imperfect communication is substantial. In the presence of message latency, every synchronous iteration is completed only after all the messages sent in the previous iteration arrive, and, therefore, the advancement of the algorithm from one iteration to the next is dependent on the longest message delay in that iteration. When messages can be lost, an agent may expect to receive a message from its neighbor while the neighbor is not aware that the message it sent did not arrive. Thus, these agents are deadlocked, each waiting for the message from the other.

In this chapter, we make the following contributions:[1]

1. We propose *Communication-Aware DCOP* (CA-DCOP),[2] an extension of the DCOP model, in which patterns of communication disturbances (e.g., message latency and message loss) can be represented.

2. We demonstrate that existing distributed local search DCOP algorithms are not robust to imperfect communication. Thus, we analyze the performance and properties of standard local search algorithms after they are adjusted to perform asynchronously in scenarios that include message latency and loss.

3. We propose an asynchronous anytime mechanism that allows any local search algorithm running in an environment with imperfect communication to report the best solution it was able to generate during its run.

[1]This chapter is based on our published papers in The 20th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2021) [29] and its extended version in the Journal of Artificial Intelligence Research (JAIR) [27].

[2]When referring to communication awareness, we do not mean that agents are aware of the communication pattern, but rather that the algorithms were designed such that they can overcome communication limitations.

4. We show that the presence of imperfect communication can have a positive impact on exploitative asynchronous local search algorithms. Our empirical results reveal that solution quality may improve as the quality of communication degrades (in terms of both message latency and message loss).

Our analysis and evaluation indicate that imperfect communication generates a discrepancy between the knowledge that agents hold and the actual state of the system. Thus, an agent may perform an action that is exploitative with respect to the information it holds, expecting to improve its state and the global state as well. However, in reality, its action degrades its state, and possibly the global state, and it unknowingly explores an unexpected part of the search space. Such explorative actions often expose agents to higher-quality solutions, allowing them to converge to better solutions.

# Chapter 4

# Latency-Aware K-Opt Monotonic Local Search for Distributed Constraint Optimization

Considering the potential negative impact of message latency on the performance of distributed algorithms, we examined a class of local-search DCOP algorithms. These algorithms are designed to provide $k$-opt guarantees, thereby ensuring that the solutions they generate cannot be improved by a group of $k$ agents [19]. One such example is MGM, a 1-opt algorithm, and its extension, MGM-2, a 2-opt algorithm. Unfortunately, their synchronous designs take advantage of the overly simplistic communication assumptions in the DCOP model that do not reflect real-world scenarios. To address these limitations, researchers introduced an asynchronous $k$-opt algorithm called Distributed Asynchronous Local Optimization (DALO) to solve DCOPs. However, as we show in this chapter, DALO's design lacks robustness in scenarios with message delays, restricting its applicability.

In this chapter, we propose Latency-Aware algorithms that guarantee 1-opt and 2-opt solutions and are robust to message latency. [1]

1. We introduce Latency-Aware Monotonic Distributed Local Search (LAMDLS), a novel local search algorithm that exhibits resilience to message latency. It is guaranteed to be monotonic and converges to a 1-opt solution, which is similar to the properties of the MGM algorithm mentioned in [15]. The algorithm utilizes an ordered coloring scheme to prevent neighboring agents from simultaneously replacing assignments and preventing agents from waiting for messages, which is a feature of MGM.

[1] This chapter is based on our novel algorithms, which have already been presented in the Journal of Artificial Intelligence Research (JAIR) [27] and are currently being developed further for submission to the 33rd International Joint Conference on Artificial Intelligence (IJCAI 2024) [28].

2. We present an extended version of LAMDLS termed LAMDLS-2. This algorithm allows agents to collaborate in pairs and coordinate the selection of their value assignments while ensuring that monotonicity is preserved, and a 2-opt solution is achieved. LAMDLS-2 facilitates sequential adjustments to the values among partnering agents. By employing a unique pairing-selection process and an ordered coloring scheme, agents can modify the values concurrently for unconstrained pairs.

3. Our work includes proofs of the theoretical properties and empirical evaluation of the proposed algorithms. Specifically, we present theoretical evidence that supports the monotonicity and convergence of the 1 and 2 opt solutions for LAMDLS and LAMDLS-2, respectively. To validate our findings, we conduct empirical evaluations in diverse environments with varying latency patterns. The performance of our algorithms was compared to that of MGM and MGM-2, and the results demonstrated that LAMDLS and LAMDLS-2 achieved convergence at significantly faster rates.

# Chapter 5

# The Effect of Asynchronous Execution and Imperfect Communication on Max-sum Belief Propagation

In Chapter 3 we investigate the effect of message latency and loss on standard distributed local search algorithms (e.g.,MGM and DSA) and show that imperfect communication has a significant positive effect on the performance of the asynchronous versions of these algorithms [27, 29]. Imperfect communication generates an exploration effect that significantly improves the quality of the solutions found. This chapter focuses on the investigation of the effect of imperfect communication on distributed incomplete inference algorithms (e.g., Max-sum), which have been shown to be very successful [2, 3].

Max-sum has been presented both as an asynchronous and as a synchronous algorithm [4, 6, 43]. To the best of our knowledge, the implications of this difference in the execution of the algorithm on its performance have not yet been studied. Moreover, when message loss is considered, the synchronous version is not applicable, because an agent may remain idle while waiting for the arrival of a message that is lost. While message latency does not affect the actions that agents perform (it only delays them) in the synchronous version, intuitively, it is expected to have a major effect on the performance of the asynchronous version. This is because the beliefs included in the messages are used by agents in the construction of beliefs that they propagate to others and in their assignment selection. In asynchronous execution, belief construction and assignment selection may be performed while considering imbalanced and inconsistent information.

In this chapter, we make the following contributions[1]:

[1]This chapter is based on our published papers in The 27th International Conference on Principles and Practice of Constraint Programming (CP 2021) [45] and its extended version in the Artificial Intelligence Journal (AIJ) [44].

1. We analyze the properties of the two execution versions of Max-sum, synchronous and asynchronous. More specifically, using backtrack cost trees [42], we investigate the possible differences between propagated beliefs in synchronous and asynchronous executions of Max-sum.

2. We investigate the effect of damping on asynchronous Max-sum. Although there are clear indications (both empirical and theoretical) that damping improves the performance of the synchronous version of the Max-sum [3, 42], to the best of our knowledge, the effect of damping on the asynchronous version of the Max-sum has not been studied before. We analyzed this effect both theoretically and empirically. Both indicate that damping reduces the differences between synchronous and asynchronous executions.

3. We investigate the performance of the different versions of the algorithm in the presence of message latency and message loss. While the beliefs propagated and the computations that agents perform are not affected by message latency in the synchronous version (only delayed), this is not true for the asynchronous version. Our empirical results reveal that damping reduces these differences. Moreover, the version of Max-sum proposed by Cohen *et al.* [3] that includes both damping and splitting maintains its fast convergence properties and the quality of solutions, even in asynchronous execution with message delays and when many messages are lost.

# Chapter 6

# Asynchronous Communication Aware Multi-Agent Task Allocation

In this chapter, we focus on task allocation problems, which present significant challenges in realistic scenarios. These challenges extend beyond the coordination of agents with degradation and unreliability of communication. Moreover, these scenarios are often highly dynamic and characterized by the emergence of new events or changes in the status of existing events [37]. The identification of dynamic events is most likely due to the agents performing in the environment. Thus, we expect agents to be able to reinitialize the solving process when necessary [7, 9].

Fisher Market Clearing Task Allocation (FMC_TA) [17] is an algorithm that was proposed for solving problems where a team of heterogeneous agents needs to cooperate in an environment that includes multiple tasks, which require ad-hoc coalitions of agents with different skills in order to properly handle them , for more details see...in the background chapter . FMC_TA was shown to dominate state-of-the-art centralized and distributed task allocation algorithms. These included general optimization algorithms such as Simulated annealing and designated algorithms such as Coalition formation with a look ahead (CFLA) [6, 17, 31].

However, similar to the observations made for DCOP algorithms in previous chapters, FMC_TA as proposed in [17] is synchronous [15, 40, 41]. Moreover, the algorithm presented in [17] is dependent on the assumption that agents hold regarding the existence and importance of tasks that need to be performed. Thus, the team of agents was not independent and relied on updates from a centralized system.

To overcome these limitations of FMC_TA when facing realistic dynamic scenarios and scenarios in which communication is imperfect, we follow the approach presented in previous chapters and make the following contributions[1].

1. We propose FMC_ATA, an asynchronous version of the algorithm that was designed while taking into consideration the possibility that messages can be delayed or lost (i.e., communication aware design). In FMC_ATA, agents perform computations upon receiving a message, considering the last message that arrives from each of their neighbors. The algorithm is performed in a single distributed asynchronous phase, in which agents determine both the allocation of tasks and their schedules. Moreover, agents in FMC_ATA can detect dynamic events, such as new tasks that need to be performed or a change in the importance of a task that is currently being handled, and trigger execution of the algorithm, to allow the team to adapt to the evolved scenario.

2. We demonstrate empirically that, not only does the asynchronous version (FMC_ATA) converge to the same solution as FMC_TA, but that it is also robust to message delays and to message loss up to some extent.

3. We further investigated the properties of scenarios in which a distributed implementation of FMC_ATA is preferred over a centralized implementation in which a central system is updated by the agents regarding dynamic events, calculates and updates allocation and updates the agents. Our results show that, in the presence of message latency, a clear threshold exists for distributed performance to be motivated, and when message loss cannot be avoided, distributed performance is always preferred.

[1]This chapter is based on our published papers in The 32nd International Joint Conference on Artificial Intelligence (IJCAI 2023) [24] and its extended version is currently being developed for submission to the Artificial Intelligence Journal (AIJ) [26].

# Chapter 7

# Summary and Conclusions

## 7.1 Conclusions

The realm of MAO faces a significant challenge in addressing the difficulties posed by imperfect communication environments, which are essential elements in representing real-world characteristics in practical applications. These difficulties include message latency and message loss, which can severely affect the performance, reliability, and applicability of distributed algorithms. Our analysis identifies major limitations in the empirical and theoretical properties of such algorithms. Current MAO methods often assume ideal communication conditions that are rarely present in real-world applications. This challenge emphasizes the need for a robust and adaptable algorithmic design that can effectively function within the constraints of imperfect communication environments. The objective of this thesis is to address this notable gap in the existing literature by introducing communication-aware models and novel algorithmic designs. This approach offers a comprehensive suite of solutions that not only advance the field of MAO but also significantly increase the applicability of MAO algorithms in the real world, where imperfect communication is an inherent challenge.

In this thesis, our exploration focus was on DCOPs and GTAPs, offering a dual perspective spanning both abstraction and practical problems. This approach was beneficial in that it enabled us to thoroughly examine our hypotheses using a controlled, abstract setting, and provided us with the opportunity to confirm our results in more practical, real-world contexts.

In Chapter 3, we explored the challenges and implications of message latency and loss on distributed local search algorithms for DCOPs and highlighted the limitations of synchronous algorithms. Our investigation revealed that asynchronous versions of DSA and MGM algorithms exhibited improved explorative effects and a faster rate of convergence. In light of these promising results, we designed an Asynchronous Anytime Mechanism to mitigate the limitations of existing anytime mechanisms, which are only applicable to synchronous algorithms, as proposed by Zivan, Okamoto, and Peled [41]. Although asynchronous versions provide improved results in the presence of imperfect communication, they do not preserve certain theoretical properties of their synchronous counterparts, necessitating the development of novel solutions, which led to the focus of Chapter 4.

In the subsequent chapter, we presented the Latency-Aware Monotonic Distributed Local Search (LAMDLS) algorithm for resolving DCOPs and ensuring convergence to 1-opt solutions. Subsequently, we extended our design to LAMDLS-2, which guarantees convergence to 2-opt solutions. Our findings reveal that our approach, based on the ordered color scheme, enables agents to compute their assignments more proactively. This reduces the coordination effort required and leads to a significantly faster convergence rate compared to MGM and MGM-2.

In this continuation of our investigation, we delve into the class of inference algorithms, focusing specifically on the Max-sum algorithm and its variants. Our analysis examines the differences between synchronous and asynchronous executions and the benefits of employing damping to mitigate the adverse effects of message latency. Through our theoretical and empirical examinations, we aimed to enhance our understanding of the Max-sum algorithm in distributed settings where communication challenges are prevalent.

The research also focused on the design of the FMC_ATA algorithm, which is a robust solution for multi-agent task allocation problems that can provide high-quality results, even under adverse communication conditions. This algorithm demonstrated a significant improvement over its predecessors in terms of real-world applicability. Additionally, we examined the performance of distributed algorithms compared to centralizing all information and performing central solving, and our findings establish the conditions under which distributed algorithms are justified.

In conclusion, this thesis has challenged the essential drawbacks posed by imperfect communication in Multi-Agent Optimization, presenting creative and adaptable solutions. By developing new algorithms, our work fills a crucial gap in the existing research and remarkably improves the practical application of MAO. Our results and analysis established a strong foundation for future advancements in MAO, aiming to inspire further research into creating more resilient and efficient systems.

## 7.2 Future Work

This thesis established a solid foundation for exploring numerous potential directions for the development of applicable MAO algorithms. Promising areas for future exploration are the examination of additional constrained communication graphs that reflect real-world challenges, such as limited bandwidth or unique network topologies, exemplified by the smart home problem identified by Rust, Picard, and Ramparany [32], where algorithms strive to reduce the communication costs among agents. The asynchronous algorithmic strategies developed in this study, designed with communication awareness in mind, may face limitations in these contexts.

Our study revealed a gap in the comprehension of the practical benefits provided by asynchronous variants of distributed algorithms in the context of imperfect communication. It has become evident that these algorithms either maintain the accuracy of the results in comparison to their synchronous counterparts or even improve them. A potential avenue for future research could involve the application of explainable AI techniques to gain a more profound understanding of the underlying reasons for the effectiveness of agents that operate with outdated information in asynchronous settings.

Furthermore, future research could focus on refining the LAMDLS and LAMDLS-2 algorithms introduced in this study. Building upon our initial contributions, we can extend our methodology to encompass a general $k$-opt algorithm. Exploration of the DALO-k algorithm, as discussed by Kiekintveld et al. [13], also presents a valuable avenue. Adapting its design to include communication-aware features could help to address some of the identified limitations.

# Bibliography


[1] M. Basharu, I. Arana, and H. Ahriz. “Solving DisCSPs with penalty driven search”. In: *Proceedings of AAAI*. 2005, pp. 47–52.

[2] Ziyu Chen et al. “A class of iterative refined Max-sum algorithms via non-consecutive value propagation strategies”. In: *Auton. Agents Multi Agent Syst.* 32.6 (2018), pp. 822–860.

[3] Liel Cohen, Rotem Galiki, and Roie Zivan. “Governing convergence of Max-sum on DCOPs through damping and splitting”. In: *Artificial Intelligence Journal (AIJ)* 279 (2020).

[4] Yanchen Deng and Bo An. “Speeding Up Incomplete GDL-based Algorithms for Multi-agent Optimization with Dense Local Utilities”. In: *Proceedings of the 29th International Joint Conference on Artificial Intelligence, (IJCAI)*. 2020, pp. 31–38.

[5] A. Farinelli et al. “Decentralised Coordination of Low-Power Embedded Devices Using the Max-Sum Algorithm”. In: *Proceedings of the International Conference on Autonomous Agents and Multiagent Systems*. 2008, pp. 639–646.

[6] A. Farinelli et al. “Decentralized Coordination of Low-Power Embedded Devices Using the Max-Sum Algorithm”. In: *Proceedings of the International Conference on Autonomous Agents and Multiagent Systems*. 2008, pp. 639–646.

[7] Alessandro Farinelli, Luca Iocchi, and Daniele Nardi. “Distributed on-line dynamic task assignment for multi-robot patrolling”. In: *Autonomous Robots* 41.6 (2017), pp. 1321–1345.

[8] F. Fioretto, E. Pontelli, and W. Yeoh. “Distributed Constraint Optimization Problems and Applications: A Survey”. In: *Journal of Artificial Intelligence Research* 61 (2018), pp. 623–698.

[9] Ferdinando Fioretto, Enrico Pontelli, and William Yeoh. “Distributed constraint optimization problems and applications: A survey”. In: *Journal of Artificial Intelligence Research* 61 (2018), pp. 623–698.

[10] Ferdinando Fioretto, William Yeoh, and Enrico Pontelli. "A Multiagent System Approach to Scheduling Devices in Smart Homes". In: *Proceedings of the 16th Conference on Autonomous Agents and MultiAgent Systems*. 2017, pp. 981–989.

[11] K. D. Hoang et al. "A Large Neighboring Search Schema for Multi-agent Optimization". In: *Proceedings of CP*. 2018, pp. 688–706.

[12] E. G. Jones, M. B. Dias, and A. Stentz. "Learning-enhanced Market-based Task Allocation for Oversubscribed Domains". In: *Intelligent Robots and Systems, 2007. IROS 2007. IEEE/RSJ International Conference on*. IEEE. San Diego, CA, 2007, pp. 2308–2313.

[13] Christopher Kiekintveld et al. "Asynchronous algorithms for approximate distributed constraint optimization with quality bounds." In: *Proceedings of the International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*. Vol. 10. 2010, pp. 133–140.

[14] N. A. Lynch. *Distributed Algorithms*. Morgan Kaufmann Series, 1997.

[15] R. Maheswaran, J. Pearce, and M. Tambe. "Distributed Algorithms for DCOP: A Graphical Game-Based Approach". In: *Proceedings of PDCS*. 2004, pp. 432–439.

[16] P. J. Modi et al. "ADOPT: Asynchronous Distributed Constraint Optimization with Quality Guarantees". In: *Artificial Intelligence* 161.1–2 (2005), pp. 149–180.

[17] Sofia Amador Nelke, Steven Okamoto, and Roie Zivan. "Market Clearing-based Dynamic Multi-agent Task Allocation". In: *ACM Transactions of Intelligent Systems Technology.* 11.1 (2020), 4:1–4:25.

[18] D. T. Nguyen et al. "Distributed Gibbs: A Linear-Space Sampling-Based DCOP Algorithm". In: *Journal of Artificial Intelligence Research* 64 (2019), pp. 705–748.

[19] J. Pearce and M. Tambe. "Quality Guarantees on k-Optimal Solutions for Distributed Constraint Optimization Problems". In: *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI)*. 2007, pp. 1446–1451.

[20] Jonathan P Pearce and Milind Tambe. "Quality Guarantees on k-Optimal Solutions for Distributed Constraint Optimization Problems". In: *Proceedings of IJCAI*. 2007, pp. 1446–1451.

[21] J. Pearl. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. San Francisco, California: Morgan Kaufmann, 1988.

[22] Arseni Pertzovskiy, Roie Zivan, and Noa Agmon. "CAMS: Collision Avoiding Max-Sum for Mobile Sensor Teams". In: *Proceedings of the 2023 International Conference on Autonomous Agents and Multiagent Systems*. 2023, pp. 104–112.

[23] A. Petcu and B. Faltings. "A Scalable Method for Multiagent Constraint Optimization". In: *Proceedings of the Nineteenth International Joint Conference on Artificial Intelligence (IJCAI)*. 2005, pp. 1413–1420.

[24] Ben Rachmut, Sofia Amador Nelke, and Roie Zivan. "Asynchronous Communication Aware Multi-Agent Task Allocation". In: *Proceedings of the Thirty-Second International Joint Conference on Artificial Intelligence (IJCAI-23)*. 2023, pp. 262–270.

[25] Ben Rachmut, Sofia Amador Nelke, and Roie Zivan. "Asynchronous Communication Aware Multi-Agent Task Allocation (EA)". In: *Proceedings of the 2023 International Conference on Autonomous Agents and Multiagent Systems*. 2023, pp. 2340–2342.

[26] Ben Rachmut, Sofia Amador Nelke, and Roie Zivan. "Asynchronous Communication Aware Multi-Agent Task Allocation[in progress]". In: *Artificial Intelligence* (2024).

[27] Ben Rachmut, Roie Zivan, and William Yeoh. "Communication-Aware Local Search for Distributed Constraint Optimization". In: *Journal of Artificial Intelligence Research* 75 (2022), pp. 637–675.

[28] Ben Rachmut, Roie Zivan, and William Yeoh. "Latency-Aware 2-Opt Monotonic Local Search for Distributed Constraint Optimization [in progress]". In: *Proceedings of the Thirty-Third International Joint Conference on Artificial Intelligence (IJCAI-24)*. 2024.

[29] Ben Rachmut, Roie Zivan, and William Yeoh. "Latency-aware local search for distributed constraint optimization". In: *Proceedings of the 2021 International Conference on Autonomous Agents and Multiagent Systems*. 2021, pp. 1019–1027.

[30] S. D. Ramchurn et al. "Coalition formation with spatial and temporal constraints". In: *Proceedings of the 9th International Conference on Autonomous Agents and Multiagent Systems (AAMAS-10)*. Toronto, Canada, 2010, pp. 1181–1188.

[31] Sarvapali D. Ramchurn et al. "Decentralized Coordination in RoboCup Rescue". In: *Computer* 53.9 (2010), pp. 1447–1461.

[32] Pierre Rust, Gauthier Picard, and Fano Ramparany. "Resilient distributed constraint reasoning to autonomously configure and adapt IoT environments". In: *ACM Transactions on Internet Technology* 22.4 (2022), pp. 1–31.

[33] A. Schoneveld, J. F. de Ronde, and P. M. A. Sloot. "On the Complexity of Task Allocation". In: *Journal of Complexity* 3 (1997), pp. 52–60.

[34] M. Smith and R. Mailler. “Getting What You Pay For: Is Exploration in Distributed Hill Climbing Really Worth It?” In: *Proceedings of IAT*. 2010, pp. 319–326.

[35] M. Vinyals et al. “Quality Guarantees for Region Optimal DCOP algorithms”. In: *Proceedings of the International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*. 2011, pp. 133–140.

[36] W.E. Walsh and M.P. Wellman. “A market protocol for decentralized task allocation”. In: *Proceedings of the International Conference on Multi-Agent Systems*. 1998, pp. 325 –332.

[37] Changyun Wei, Koen V Hindriks, and Catholijn M Jonker. “Dynamic task allocation for multi-robot search and retrieval tasks”. In: *Applied Intelligence* 45.2 (2016), pp. 383–401.

[38] Chen Yanover, Talya Meltzer, and Yair Weiss. “Linear Programming Relaxations and Belief Propagation - An Empirical Study”. In: *Journal of Machine Learning Research* 7 (2006), pp. 1887–1907.

[39] M. Yokoo and K. Hirayama. “Distributed breakout algorithm for solving distributed constraint satisfaction problems”. In: *Proceedings of the International Conference on Autonomous Agents and Multiagent Systems*. 1996.

[40] W. Zhang et al. “Distributed Stochastic Search and Distributed Breakout: Properties, Comparison and Applications to Constraint Optimization Problems in Sensor Networks”. In: *Artificial Intelligence* 161.1–2 (2005), pp. 55–87.

[41] R. Zivan, S. Okamoto, and H. Peled. “Explorative anytime local search for distributed constraint optimization”. In: *Artificial Intelligence* 211 (2014).

[42] Roie Zivan, Omer Lev, and Rotem Galiki. “Beyond Trees: Analysis and Convergence of Belief Propagation in Graphs with Multiple Cycles”. In: *Proceedings of the 34th International Conference of the Association for the Advancement of Artificial Intelligence (AAAI)*. 2020, pp. 7333–7340.

[43] Roie Zivan et al. “Balancing exploration and exploitation in incomplete Min/Max-sum inference for distributed constraint optimization”. In: *Journal of Autonomous Agents and Multi-Agent Systems (JAAMAS)* 31.5 (2017), pp. 1165–1207.

[44] Roie Zivan et al. “Effect of asynchronous execution and imperfect communication on max-sum belief propagation”. In: *Autonomous Agents and Multi-Agent Systems* 37.2 (2023). ISSN: 1387-2532.

[45] Roie Zivan et al. "The effect of asynchronous execution and message latency on Max-Sum". In: *27th International Conference on Principles and Practice of Constraint Programming (CP 2021)*. Schloss Dagstuhl-Leibniz-Zentrum für Informatik. 2021.