STRATEGIC PATH OPTIMIZATION IN PARTIALLY KNOWN ENVIRONMENTS

by

Adam Hudson

A thesis submitted to the faculty of
The University of North Carolina at Charlotte
in partial fulfillment of the requirements
for the degree of Master of Science in
Electrical Engineering

Charlotte

2024

Approved by:

______________________________
Dr. Dipankar Maity

______________________________
Dr. Jim Conrad

______________________________
Dr. Srinivas Akella

ABSTRACT

ADAM HUDSON. Strategic Path Optimization in Partially Known Environments.
(Under the direction of DR. DIPANKAR MAITY)

This thesis addresses the challenge of autonomous navigation in partially known environments, focusing on mobile robots equipped with limited sensing capabilities. Path planning and sensor usage are closely linked in such environments, as each sensor activation incurs a cost. Thus, the robot faces a trade-off: it can either follow longer, sensor-free paths or activate sensors to take shorter, more efficient routes. To address this, we develop two joint decision-making frameworks integrating path planning with strategic sensor activations to optimize navigation efficiency under resource constraints.

A *Regret-aware Joint Sensing and Path Planning* presents a combined sensing and control framework designed to minimize path length while efficiently using sensing. This methodology employs a joint cost map to evaluate the *value of the information* gained from each potential sensor activation, prioritizing sensing only where substantial information gain or path improvement is likely. Initial results demonstrate that this approach allows the robot to allocate sensors effectively, avoiding unnecessary activations and improving navigation efficiency compared to simpler controllers.

A *Sensor-aware Planner and Regret-based Cost Function* method builds on the first by expanding the robot's planning capability to incorporate sensor use as part of the path planner's state space. This methodology includes a modified A* algorithm that operates within an enlarged, multi-dimensional search space, representing the environment and the robot's real-time available sensing budget. An edge-based cost function dynamically evaluates local decisions based on *regret* and *information value*. The proposed sensor-aware A* planner enables the robot to anticipate future sensor activations, strategically utilizing the sensor across its path. Repeated simulation shows statistical evidence that this approach enhances the robot's ability to prioritize critical sensing locations, improving path efficiency while reducing overall sensor usage.

Together, these methodologies advance the robot's capacity to navigate unknown environments by balancing short-term and long-term sensing strategies, providing a scalable framework for resource-limited exploration in complex, partially mapped environments.

# DEDICATION

To my family, whose unwavering support and encouragement have guided me every step of the way.

To my mentors throughout my life, who have challenged and inspired me to push beyond my limits.

And to all those who continue to explore the unknown, driven by curiosity and the pursuit of knowledge, paving the way for the next generation of innovators.

# TABLE OF CONTENTS

# LIST OF TABLES

# LIST OF FIGURES

# LIST OF SYMBOLS

| | |
|---|---|
| $\mathbf{p}$ | Path within free cells |
| $\mathcal{F}$ | Set of all free cells |
| $\mathcal{I}_t$ | Robot's Information at time $t$ |
| $\mathcal{O}$ | Set of all occupied cells |
| $\theta_t$ | Sensor activation decision at time $t$ |
| $\mathcal{U}$ | Set of all unknown cells |
| $\mathcal{W}$ | World |
| $a_t$ | Action at time $t$ |
| $c(\cdot)$ | cost function |
| $g$ | Goal position |
| $H$ | Path Length |
| $N$ | Number of sensor that can be used |
| $N_t$ | Sensors used at time $t$ |
| $P$ | Set of all paths $p$ |
| $Q$ | Set of all paths $q$ |
| $q$ | Optimistic path through free and unknown cells |
| $r_S$ | sensing radius of sensor |
| $s$ | Starting position |
| $S(\cdot)$ | Sensor activation function |

| | |
|---|---|
| $t$ | Time |
| $x_t$ | Robot's position at time $t$ |

# CHAPTER 1: INTRODUCTION

*On Mars, every movement counts.* Autonomous robots exploring uncharted terrains must make critical decisions with every step, often with limited knowledge of what lies ahead [7, 8]. Unlike Earth-based systems, which can rely on rapid human oversight, Mars rovers must operate independently in a landscape where communication delays stretch for minutes and resources are scarce [9]. Each movement and sensor activation brings them closer to their destination but also closer to depleting their power reserves. The challenge becomes clear: **how can a robot efficiently explore the unknown, gathering just enough information to make intelligent decisions while conserving the resources required to reach its destination?**

Navigating a partially known environment with limited information and resource constraints is a crucial task for autonomous systems [10, 11], not only in space exploration but also in diverse real-world applications, from disaster response to environmental monitoring. In these scenarios, the robot's ability to balance exploration (to reveal unknowns) with efficiency (to conserve resources) defines its success. This thesis dives into this problem, proposing solutions for how an autonomous robot can manage sensor usage and path planning in partially known and resource-limited environments.

## 1.1 Problem Statement and Research Scope

The core focus of this research is developing an efficient joint perception and planning algorithm for a partially known environment with explicit sensing constraints. This thesis addresses the challenge of optimizing path planning under conditions of uncertainty and limited sensor availability. An autonomous robot, tasked with moving from a defined starting point to a goal, must make strategic decisions about when and where to activate its sensors, revealing unknown regions only when necessary.

The aim is to develop an algorithm that balances two priorities. First, Perception: using sensors selectively to gather crucial information about unknown areas. Second, Planning: ensuring efficient, goal-directed movement to conserve resources. This balance is particularly challenging due to the robot's partial knowledge of the environment and its limited sensor activations. Navigating under these constraints requires a control system that dynamically adjusts its sensing and movement strategies based on the operating environment, resource availability, and mission objectives.

## 1.2 Research Motivation and Real-World Relevance

The significance of this research extends beyond theoretical exploration, encompassing practical applications in diverse and challenging environments. Mars exploration serves as a prime example, where autonomous robots face strict limitations on sensor usage due to power and time constraints. These constraints mirror challenges on Earth, such as disaster response scenarios, where robots must navigate through hazardous areas with finite resources while balancing urgent mission objectives. Military operations further highlight the relevance of this work, as autonomous systems are increasingly deployed in contested environments where sensor activations must be carefully managed to avoid detection, conserve energy, and maintain operational efficiency.

Figure 1.1 illustrates the breadth of this research's applications and its potential extensions. The applications range from off-road navigation and space exploration to search-and-rescue missions, all of which require precise management of sensing and navigation resources. Meanwhile, the extensions include innovative advancements such as incorporating rechargeable resources (e.g., solar panels or radiothermal generators), adapting to dynamic obstacles in the environment, and scaling to multi-agent systems for coordinated exploration and decision-making.

No matter the application–be it Mars exploration, disaster response, or military navigation–an autonomous agent will always operate under constraints such as limited sensors, finite power, or strict time budgets. The strategies proposed in this research directly address these challenges by integrating sensor-aware path planning with real-time navigation, enabling robots to adapt to dynamic and uncertain environments while maintaining mission objectives. By grounding these strategies in practical scenarios, this work lays the foundation for advancements in space exploration, search-and-rescue operations, and military applications, ensuring that robots can operate effectively in even the most demanding and resource-constrained environments.

## 1.3 Research Objectives

The objectives of this thesis align with the goal of creating a joint perception and planning algorithm capable of navigating partially known environments efficiently. The specific objectives are as follows:

1. To develop a control system that dynamically balances exploration and goal-directed movement, optimizing both *long-term gains* and *short-term decisions* to minimize overall path length. The system aims to prioritize efficient navigation by adjusting its approach based on the terrain and resource availability.

2. To implement an adaptive balance between sensing and control, allowing the system to determine when sensor activation is necessary based on the robot's position, environmental uncertainty, and resource constraints. This dynamic approach will enable the robot to adapt to varying levels of uncertainty and respond effectively to the need for information.

3. To evaluate the algorithm's performance across environments with different levels of uncertainty. This evaluation will assess path length, sensor usage, and adaptability to ensure the algorithm is robust and efficient under diverse condi-

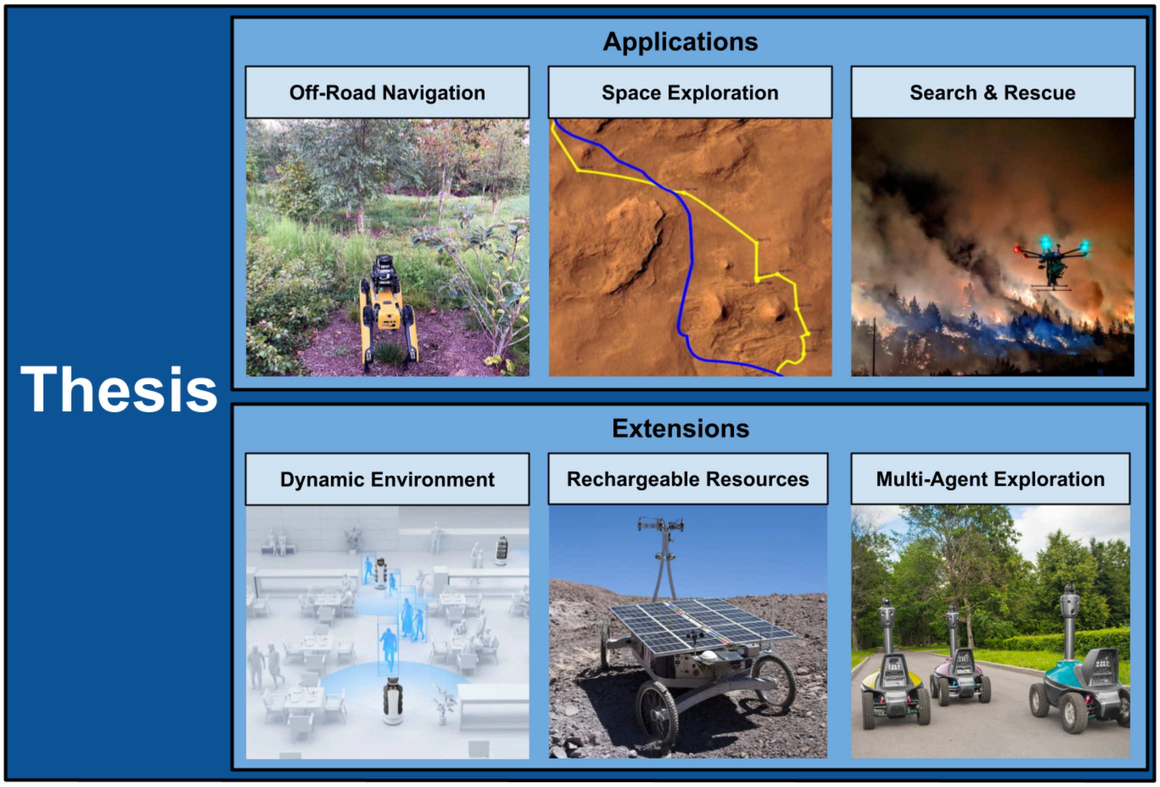



Figure 1.1: Applications and extensions of the thesis. Application image sources [1, 2, 3]. Extension image sources [4, 5, 6].

tions. By analyzing performance across these metrics, the research will provide insights into the tradeoffs between exploration and efficiency in practical navigation tasks.

## 1.4 Thesis Contribution

This thesis contributes to the field of autonomous robotic navigation by introducing solutions that address the challenges of resource-constrained exploration in partially known environments. The key contributions, depicted in Figure 1.1, outline three major advancements and their potential applications, alongside promising extensions for future exploration.

1. *Joint Perception and Planning Algorithm*: A novel algorithm that balances perception and planning, enabling efficient navigation in environments with explicit constraints on sensor usage and energy consumption. This algorithm is

designed to handle uncertainty and resource limitations effectively, providing a method for robots to achieve mission goals with minimal resource expenditure.

2. *Dynamic Control Approach*: This thesis demonstrates a control approach that optimizes decision-making in both the short and long term. By dynamically adjusting the robot's behavior based on immediate conditions, the control system ensures efficient, adaptable navigation even in challenging, resource-limited scenarios.

3. *Abstract Simulation Environment*: An abstracted simulation replicates real-world terrains, including forests, rough landscapes, and extraterrestrial surfaces. This environment serves as a flexible testing platform for autonomous navigation strategies, allowing for rigorous evaluation under conditions that closely mirror practical challenges.

The contributions of this thesis lay a robust foundation for developing autonomous systems capable of balancing exploration and efficiency under constraints, advancing the potential for real-world applications such as off-road navigation, space exploration, and search and rescue. By enabling autonomous systems to operate effectively in diverse and demanding environments, this research also sets the stage for impactful extensions, including incorporating rechargeable resources from power sources like solar panels and radioisotope thermoelectric generators, adapting to dynamic obstacles, and leveraging multi-agent systems for coordinated exploration. Together, these advancements highlight the versatility and transformative potential of the presented methodologies, paving the way for autonomous exploration in challenging settings like Mars, dense forests, or disaster zones.

## 1.5 Thesis Structure

The thesis is organized as follows:

Chapter 2: Literature Review - This chapter surveys existing work in autonomous navigation and sensor usage, covering historical approaches, modern methodologies, and a critical analysis of challenges and gaps in the field.

Chapter 3: Problem Formulation - A detailed definition of the problem, including core concepts, assumptions, and the formal problem statement. This chapter clarifies the fundamental components and constraints that shape the research.

Chapter 4: Preliminary Work - Description of initial experiments and exploratory work, including the development of basic strategies for navigating partially known environments.

Chapter 5: Methodology 1: Regret-aware Joint Sensing and Path Planning- This chapter presents the joint sensing and control framework, detailing how sensor usage and path planning integrate to achieve optimal navigation. *(Paper under review)*

Chapter 6: Methodology 2: Sensor-aware Planner and Regret-based Cost Function- Introduction of a global-local planning approach that updates dynamically based on new environmental information. *(Paper in progress)*

Chapter 7: Overall Results and Analysis - Comparative analysis of both methodologies, evaluating tradeoffs between sensing and control across varied scenarios.

# CHAPTER 2: Literature Review

Navigating partially known environments while balancing sensing and control requires innovative strategies that adapt to uncertainty. Prior research spans themes such as reinforcement learning, active sensing[12], value of information (VOI) [13], and high-dimensional path planning. These efforts have contributed foundational theories, practical methodologies, and key insights, but a conclusive, unified solution remains elusive. This review synthesizes these contributions to establish the state of the art, highlight gaps, and position the existing work.

## 2.1 Foundations of Sensing and Planning

Historical advances in sensing and planning algorithms have provided the theoretical backbone for modern approaches. Dijkstra's algorithm [14] laid the foundation for shortest-path planning, offering a deterministic solution that inspired heuristic methods like A* [15]. Stentz [16] extended these foundational ideas to partially known environments with the introduction of the D* algorithm. D* efficiently replans paths in dynamic and uncertain environments, addressing the limitations of global replanning and local reactive adjustments. This advancement marked a significant milestone in enabling optimal and computationally efficient navigation for exploratory robots.

Bajcsy [12] introduced active perception, emphasizing the interplay between control strategies and information gain–a concept that underpins much of modern active sensing research. Similarly, Howard's information value theory [17] formalized the relationship between probabilistic uncertainty and economic decision-making, while Khatib [18] pioneered artificial potential fields for real-time obstacle avoidance. Together, these methods laid the groundwork for integrating sensing and planning in autonomous systems.

Dynamic environments further highlighted the need for efficient replanning algorithms. Koenig and Likhachev [19] optimized D* Lite for real-time applications, preserving D*'s optimality while improving computational performance. Chadhurbala et al.'s comparative study [20] evaluated A*, D*, and D* Lite, demonstrating their applicability to grid-based path planning. These algorithms collectively illustrate the evolution from static to dynamic planning solutions, emphasizing adaptability to partial or changing knowledge.

In the controls community, the duality between estimation and control (analogous to sensing and planning in robotics) has been well established [21, 22]. Todorov [21] demonstrated that certain classes of problems reveal a direct duality between optimal estimation and control in the traditional optimization sense. Pearson's earlier work [22] further explored this equivalence, showing how optimal strategies in one domain inform solutions in the other.

Joint sensing and control problems are often formulated as single optimization problems, incorporating both sensing and control costs [23, 24]. Baras and Bensoussan [23] and Maity and Baras [25] highlighted how sensing impacts the controller's information structure and how control actions influence the state, requiring further sensing to estimate new states. This interaction mirrors the joint sensing and planning framework envisioned for robotics.

These works provide critical theoretical insights into combining estimation and control, influencing modern strategies for integrated sensing and planning. By extending these ideas, our goal is to develop a planning framework that incorporates both sensing constraints and path-planning objectives into a unified cost map.

## 2.2 Reinforcement Learning and Policy-Based Planning

Reinforcement learning (RL) and policy-based planning approaches have been pivotal in addressing the complexities of sensing, decision-making, and control in robotics, particularly under uncertainty. These methods often rely on policies to balance trade-offs, leveraging POMDPs (Partially Observable Markov Decision Processes) to model partially observable environments. Zubek and Dietterich [26] introduced a POMDP-based framework to optimize sensor activation decisions, demonstrating the effectiveness of policies in navigating resource constraints. Bonet and Geffner [27] further developed sensor-aware planning through heuristic search in belief spaces, offering a scalable alternative to cost maps.

Chrisman and Simmons [28] explored static sensing policies, optimizing sensing operations by weighing their cost against expected utility. Cassandra et al. [29] extended this paradigm by applying discrete Bayesian models to navigation tasks, showcasing heuristic control strategies for partially observable environments. Lim [30] built upon these ideas by integrating POMDP algorithms with modern machine learning techniques, enabling visual navigation tasks. Langley [31] demonstrated how RL algorithms could selectively reduce sensory load while maintaining control, focusing on efficient coupling of sensing and actions.

RL also excels in optimizing exploration and path planning. You and Wu [32] proposed Geometric RL, which balances exploration and path efficiency by dynamically updating reward matrices. Stein et al. [33] introduced subgoal-based navigation to reduce cost-to-go metrics, while Uppal et al. [34] developed SPIN, which integrates active visual perception with navigation for mobile manipulation in unstructured environments. Kiran et al. [35] provided a comprehensive survey on deep RL techniques for autonomous driving, emphasizing the integration of behavior modeling and trajectory optimization to enhance policy efficacy. These approaches highlight RL's flexibility in adapting policies to diverse challenges, including structured and dynamic scenarios.

Applications of RL in real-world settings further underscore its utility. Delmerico et al. [36] integrated active aerial exploration with ground robot path planning, minimizing response time in search-and-rescue missions. Similarly, Wang et al. [37] applied deep RL to aerial robots for environmental monitoring, showcasing its scalability in high-dimensional planning problems. Additionally, Lluvia et al. [38] surveyed active mapping techniques, bridging RL and SLAM (Simultaneous Localization and Mapping) strategies to enable autonomous exploration. These works emphasize RL's capability to navigate complex decision spaces where sensing and control are intertwined.

While RL and policy-based methods offer practical solutions, they often lack the mathematical rigor required for generalization. Your work complements these approaches by introducing structured cost maps to formalize trade-offs between sensing and control. This integration bridges the gap between empirical RL approaches and optimization-based methodologies, offering a robust framework for navigating partially known environments.

## 2.3 Active Sensing and Perception

Active sensing focuses on dynamically acquiring information to reduce uncertainty and improve decision-making. Bajcsy [12] established active perception as a cornerstone of this field, highlighting the need to integrate sensing with control strategies. This seminal work laid the foundation for modern active sensing approaches that aim to balance sensing costs with the value of information (VoI) gained. For example, Lu et al. [39] used information potential functions to maximize sensing efficiency during navigation tasks, aligning sensing with areas of high information gain. Similarly, Wang et al. [37] demonstrated how mobile robots could optimize data collection by considering deployment time and historical sensing data to guide measurements.

Recent advancements have extended active sensing to multi-agent systems, dynamic exploration, and cost-aware strategies. Kim et al. [40] developed a cooperative strategy for quadrotor teams to explore unknown environments, optimizing coverage and efficiency through a selective graph exploration framework. Manjanna and Dudek

[41] proposed an adaptive approach for marine robots to cover spatial fields, focusing on multi-scale paths to balance information acquisition with energy costs. Lluvia et al. [38] conducted a comprehensive survey of active mapping techniques, emphasizing the integration of simultaneous localization and mapping (SLAM) with exploration for autonomous navigation. These examples showcase the diversity of active sensing methods, from single-agent optimization to collaborative multi-agent exploration, underscoring its role in navigating complex, resource-constrained environments.

Cost-effective approaches to sensing have also emerged, addressing challenges in resource management during decision-making. Hansen [42] introduced a dynamic sensing strategy using Markov decision theory, balancing sensing costs and planning benefits. This strategy dynamically adjusts sensing intervals based on uncertainty, aligning closely with modern active sensing paradigms. Similarly, Maity [24] explored optimal intermittent sensing for pursuit-evasion games, where constrained sensing budgets necessitated equilibrium strategies for effective sensing and motion. These works highlight the significance of cost-aware strategies in guiding sensing decisions under uncertainty.

In search-and-rescue scenarios, active sensing enables rapid and efficient navigation. Delmerico et al. [36] introduced an active aerial exploration framework that combines mapping with terrain classification to guide ground robots through unknown terrains. By integrating 3D reconstruction and on-the-spot learning, this approach optimizes response times while ensuring robust path planning. This emphasis on real-time decision-making aligns with Todorov's [21] concept of duality between estimation and control, where sensing (estimation) and planning (control) are treated as complementary processes.

Langley [31] explored selective sensing through statistical learning, enabling systems to reduce sensory loads while maintaining control. This notion of integrating learning with active perception resonates with broader efforts to optimize sensing and decision-making in robotics. For instance, Todorov's insights into linear-quadratic Gaussian (LQG) control provide a theoretical framework for linking sensing and planning under uncertainty, further underscoring the value of unified approaches to active perception.

Active sensing continues to evolve, with applications ranging from data-driven adaptive sampling [41] to cooperative UAV exploration [43]. These works collectively highlight the potential of active perception to drive efficient decision-making in partially known and dynamic environments. By building on these principles, your approach to integrating cost maps and sensor activation constraints provides a novel lens for understanding and optimizing the interplay between sensing and planning.

## 2.4 Value of Information (VoI) and Regret-Based Planning

Value of Information (VoI) and regret-based methodologies play crucial roles in decision-making under uncertainty, particularly in robotics and autonomous systems operating in incomplete environments. VoI quantifies the utility of additional information to reduce uncertainty, while regret measures the difference between the cost of an actual plan and the optimal plan that would have been possible with perfect knowledge. These complementary approaches help balance the trade-offs between exploration and exploitation, often critical in sensor-based path planning.

The theoretical foundations for VoI can be traced back to Howard [17], who formalized the interplay between probabilistic uncertainties and economic impacts in decision-making. Low et al. [44] extended this theory to robotic exploration by formulating adaptive path planning as a reward-maximization problem. Their information-theoretic framework emphasizes efficiency by prioritizing observations that maximize information gain while minimizing computational overhead, a critical requirement for large-scale environmental sensing.

Practical applications of VoI often involve planning under uncertainty with constraints on computational resources or sensing capabilities. Spaan et al. [45] developed a POMDP-based framework for active cooperative perception, integrating information rewards into planning to achieve specific belief-state goals. This approach demonstrated the feasibility of using VoI to drive optimal policies in collaborative scenarios, such as robot-assisted surveillance. Similarly, Likhachev and Stentz [46] proposed an information-driven approach to path clearance, where scout robots gather data on potential hazards to inform a primary robot's navigation. These methods underscore VoI's versatility in optimizing sensing and planning in diverse robotic applications.

In regret-based planning, Zhao et al. [47] introduced a Linear Temporal Logic (LTL) framework to minimize regret in partially known environments, allowing robots to balance exploration and exploitation. Building on this work, Zhao et al. [48] proposed an autonomous exploration planner for UAVs, focusing on efficient mapping in unknown environments. By dynamically adapting yaw trajectories and leveraging frontier exploration sequences, this method reduces back-and-forth maneuvers and optimizes path planning. These advancements highlight how regret minimization strategies can enhance exploration efficiency, particularly in resource-constrained scenarios.

Recent studies also explore the integration of VoI and regret to address multi-objective planning challenges. For instance, Likhachev et al. [49] developed a multi-objective variant of D* Lite, leveraging incremental graph search to optimize metrics like risk, cost, and time. Such approaches demonstrate the potential for combining VoI and regret to improve decision-making frameworks, ultimately enabling robots to operate more effectively in uncertain and resource-limited environments.

## 2.5 Sensor Constraints and Budget-Aware Planning

Resource limitations play a critical role in shaping effective sensing and planning strategies. Planetary exploration, for example, requires careful management of limited computational resources, energy, and sensor capabilities. Gerdes et al. [11] illustrate this challenge with a navigation system for Mars rovers, where efficient sensing and hazard detection allow for longer traverses in uncertain terrains. By leveraging lightweight sensing architectures and stereo vision systems, their approach demonstrates how sensing constraints can be addressed to achieve practical autonomy in resource-limited environments. Similarly, Tunstel and Howard [50] discuss sensing and perception challenges faced by planetary rovers, emphasizing the need for sensor solutions tailored to strict mass, power, and operability constraints. These examples underscore the importance of incorporating sensor budgets into planning frameworks to ensure feasibility and robustness in real-world applications.

Several studies propose frameworks for optimizing sensing while minimizing resource expenditure. Tzoumas et al. [51] address this issue through sensing-constrained Linear-Quadratic-Gaussian (LQG) control, which integrates sensor selection, estimation, and control. Their scalable algorithm selects the most critical sensors to optimize system performance, demonstrating how constraints on sensing resources can be balanced with control objectives. Similarly, Singh et al. [52] use submodular optimization to plan informative paths for multi-robot systems under energy and time budgets. Hansen [42] takes a cost-effective approach, dynamically adjusting sensing intervals based on uncertainty and cost. These results emphasize the trade-offs between information gain and resource consumption, which are central to budget-aware planning.

In contexts where sensing is intermittent or costly, strategies must adapt dynamically to maximize efficiency. Maity [24] explores optimal intermittent sensing strategies in pursuit-evasion games, demonstrating how limited sensing budgets impact a player's decision-making and overall performance. Aggarwal et al. [53] expand on this concept by introducing threshold-based policies for differential games, optimizing performance under intermittent sensing constraints. Their work, alongside Krishnamurthy's [54] game-theoretic strategies for decentralized sensor activation, provides valuable insights into balancing sensing and resource constraints in complex, dynamic environments.

Adaptivity is another critical aspect of sensor-constrained planning. Johnsen and Levorato [10] present NaviSlim, a neural navigation model that dynamically adjusts sensing and computational complexity based on the current environment. This framework demonstrates how real-time adaptability can optimize performance in constrained scenarios, such as micro-drone navigation. Wang et al. [37] extend this idea to active sensing for autonomous mobile robots, where path planning accounts for limited deployment time and measurement value, ensuring efficient sensing while respecting resource constraints. Both studies highlight the potential of adaptive models to enhance sensing efficiency without compromising mission objectives.

These methods highlight the importance of modeling sensing budgets and resource constraints in planning and control frameworks. However, fully integrating such models with adaptive decision-making systems remains challenging, particularly in dynamic and unpredictable environments. By advancing the interplay between sensing constraints and planning, these works provide a foundation for building robust, efficient autonomous systems that operate effectively within resource-limited contexts.

## 2.6 High-Dimensional and Probabilistic Path Planning

High-dimensional and probabilistic path planning addresses challenges inherent in navigating complex and uncertain environments. Foundational methods such as Probabilistic Roadmaps (PRMs) introduced by Kavraki et al. [55] and Rapidly-Exploring Random Trees (RRTs) by LaValle [56] revolutionized planning in high-dimensional spaces by employing sampling-based techniques. PRMs sample configurations to construct a graph representation of free space, enabling efficient multi-query planning. RRTs complement this by rapidly expanding towards unexplored regions, excelling in single-query scenarios. Together, these algorithms provide a basis for planning in domains where the configuration space is vast and analytically intractable.

Extensions of these methods often incorporate probabilistic reasoning to manage environmental uncertainty. Singh et al. [52] and Low et al. [57] utilized Gaussian Processes to model uncertain fields and optimize paths that maximize information gain. Singh's eSIP algorithm demonstrates submodular optimization for multi-robot scenarios, balancing efficiency and sensing cost. Low's iMASP framework redefines planning under uncertainty by optimizing information-theoretic reward functions, highlighting the potential of probabilistic models for large-scale exploration. Banfi et al. [58] expanded on these principles by incorporating hypothesis paths within probabilistic occupancy grids, dynamically refining plans as new data is acquired. These contributions bridge theoretical approaches with practical applications, emphasizing the versatility of probabilistic frameworks.

Recent advancements explore decision-making frameworks that integrate probabilistic insights with adaptive strategies. Ren et al. [49] extended the D* Lite algorithm to handle multi-objective path planning, balancing competing criteria like risk and travel time while maintaining computational efficiency through Pareto-optimal pruning. Guzzi et al. [59] analyzed the Canadian Traveller Problem to optimize navigation under uncertain edge traversability, while Manjanna et al. [41] employed adaptive sampling to efficiently cover spatial fields, showcasing applications of probabilistic modeling. Conner et al. [60] introduced a novel approach to composing local potential functions for global navigation, demonstrating the importance of blending localized strategies with overarching objectives to achieve robust solutions in constrained environments.

Collaborative and hierarchical planning strategies further enhance the capabilities of probabilistic approaches. Panov and Yakovlev [61] proposed a behavior-based model for coalition planning, where multiple agents dynamically modify their environment to achieve shared objectives. This framework leverages hierarchical state-space exploration and constrained path planning to ensure feasibility under diverse conditions. These strategies, alongside innovations in adaptive sensory feedback, underline the significance of integrating probabilistic reasoning with flexible decision-making frameworks to tackle the challenges of high-dimensional and uncertain environments effectively.

## 2.7 Similar Problems

Research addressing related challenges provides complementary insights into sensing and path planning under uncertainty. Bonet and Geffner [27] formulated planning with incomplete information as heuristic search in belief space, leveraging probabilistic reasoning to navigate unknown environments. While their approach effectively balances probabilistic actions and sensing, it does not incorporate explicit sensor budget constraints, which are critical in resource-limited scenarios. Similarly, Singh et al. [52] employed Gaussian Process-based submodular optimization to maximize information gain in multi-robot systems, and Kim et al. [40] proposed a cooperative graph-based exploration framework for UAVs, optimizing coverage in complex environments. These works focus on efficiency but lack explicit mechanisms to balance sensing costs against long-term navigation goals.

Several works highlight adaptive and reinforcement learning (RL)-based strategies for sensing and planning. Langley [31] introduced statistical learning techniques to selectively reduce sensory load while maintaining control, offering insights into coupling sensing efficiency with control actions. You and Wu [32] proposed a Geometric RL framework that balances field reconstruction and path planning through real-time updates of reward matrices. Both approaches showcase the adaptability of RL in optimizing exploration, but they do not address multi-objective scenarios where sensing budgets and regret-based planning play significant roles.

Other researchers emphasize uncertainty-aware and multi-objective approaches. Banfi et al. [58] developed an uncertainty-aware planner using probabilistic occupancy grids, dynamically refining paths based on new sensory data. However, their method does not incorporate sensing budgets or regret trade-offs, limiting its applicability in highly resource-constrained environments. Ren et al. [49] extended D* Lite to handle multi-objective optimization, balancing risk, arrival time, and other factors, but their approach focuses on Pareto fronts rather than integrating dynamic sensing updates into path planning.

Lim [30] explored POMDP-based sequential decision-making, combining theoretical guarantees with machine learning techniques for real-world applications. Spaan et al. [45] expanded POMDP frameworks to reward information gain in cooperative perception scenarios, bridging planning and active sensing. These works provide robust theoretical foundations but often overlook the computational overhead and practical constraints of sensor-limited environments.

## 2.8 Gaps in the Literature

Despite notable advancements, key challenges persist in bridging theoretical and practical approaches to sensing and planning under uncertainty. Many reinforcement learning (RL)-based methods, such as Lim's exploration of POMDP solvers [30], leverage neural networks to implicitly address partially observable environments but lack interpretability. As Lim highlights, "many deep reinforcement learning algorithms are implicitly solving a POMDP...by learning a neural network representation," which complicates their application in safety-critical domains where explainable decision-making is essential. The absence of mathematically grounded or explicitly explainable solutions limits broader adoption, especially in fields like planetary exploration or autonomous navigation.

Additionally, the integration of sensing budgets into existing frameworks remains underdeveloped. While methods like Banfi et al.'s uncertainty-aware planner [58] and Bonet and Geffner's belief-space planning [27] excel in probabilistic reasoning, they fail to explicitly incorporate resource constraints, such as limited sensor usage. Furthermore, most approaches overlook the potential synergy between value of information (VoI) and regret-based principles, which could enhance adaptability by quantifying the trade-offs between immediate exploration and long-term performance. Addressing these gaps requires solutions that are not only explainable and budget-aware but also integrate VoI and regret to create robust, resource-efficient strategies for real-world autonomous systems.

# CHAPTER 3: Problem Formulation

This section introduces the foundational concepts and assumptions that underpin this research, defining the environment, sensor behaviors, and robot capabilities. These elements are essential for understanding how the robot navigates partially known environments while balancing limited resources.

## 3.1 Core Concepts

This research operates under a set of assumptions that define the robot's environment, movement capabilities, and sensor characteristics. These assumptions ensure that the robot has a fallback strategy, even in highly uncertain scenarios, by guaranteeing that a safe but potentially less efficient path is always available.

### 3.1.1 World Definition

The environment, denoted as grid world $\mathcal{W}$, is composed of a grid of cells classified into three distinct categories: free ($\mathcal{F}$), obstacle ($\mathcal{O}$), and unknown ($\mathcal{U}$). This classification helps guide the robot's path planning and sensing decisions by defining different levels of certainty and risk. The layout of this world is illustrated in Figure 3.1.

Free cells ($\mathcal{F}$) are known traversable cells in the grid that the robot can move through safely without requiring additional sensor input. These cells are already mapped as accessible and pose no risk in terms of obstacles.

Obstacle cells ($\mathcal{O}$) are cells the robot has identified as untraversable and must avoid during navigation. Attempting to move through an obstacle cell results in failure, making it essential for the robot to exclude them from its planned route.

Unknown cells ($\mathcal{U}$) represent cells where the robot has no prior knowledge of the

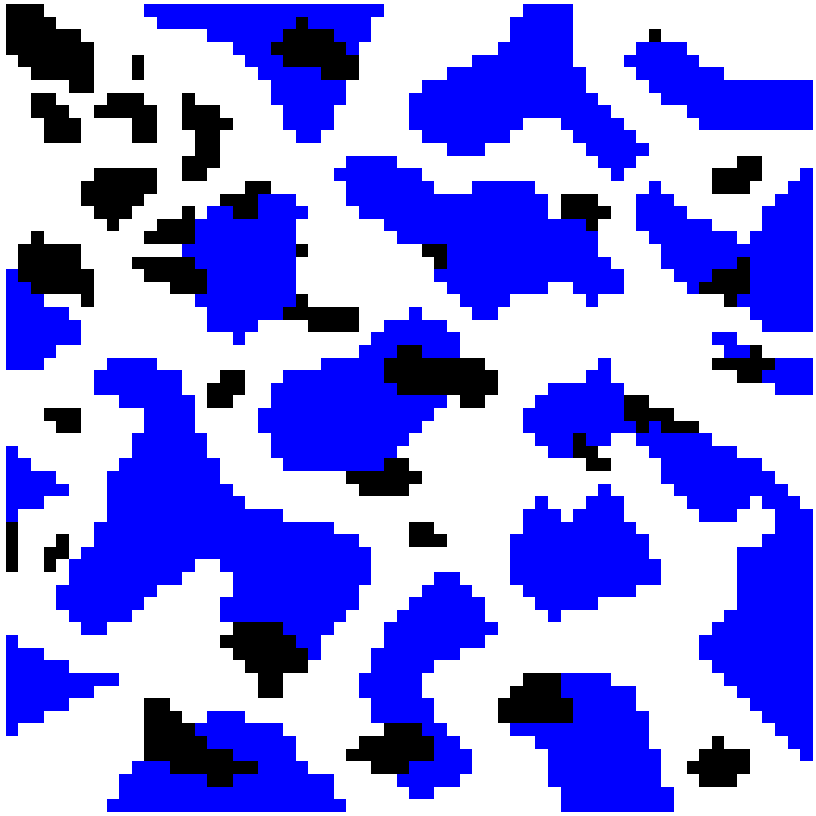


Figure 3.1: An example world. $\mathcal{O}$ is shown in black. $\mathcal{U}$ is shown in blue. $\mathcal{F}$ is shown in white.

status. These cells could be either free or obstacle, and their true nature remains unknown until the robot activates its sensor. Unknown cells are areas of high uncertainty and risk but may also offer the potential benefit of shorter paths if they turn out to be free.

The world $\mathcal{W}$ is designed so that as the robot navigates and gathers information, cell classifications can dynamically change from unknown to either free or obstacle. This reclassification allows the robot to update its path-planning approach as it collects data about its environment, enhancing its ability to navigate effectively toward the goal.

### 3.1.2 Robot Definition

The robot operates by moving through $\mathcal{W}$ and gathering information to resolve unknown cells. It is equipped with a sensor that can be selectively activated to reveal

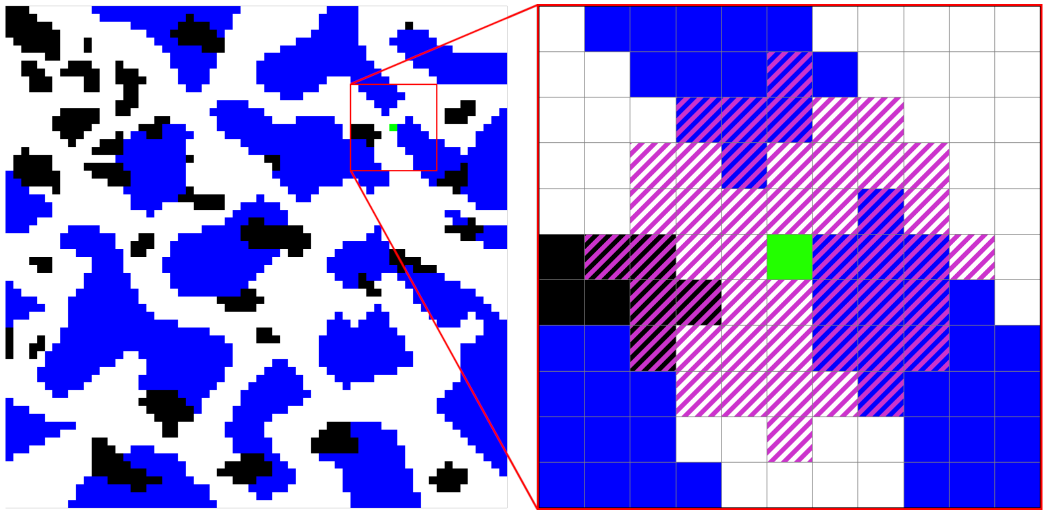


Figure 3.2: Sensor operation being demonstrated. The green point is cell $x_t$, the origin of the sensing action.

details about cells within a specific sensing region. At any given time $t$, the robot can execute one of four movement actions, denoted $a_t \in \{\texttt{up}, \texttt{down}, \texttt{left}, \texttt{right}\}$, each allowing it to traverse to an adjacent cell in the specified direction. Movement is restricted to cells classified as free ($\mathcal{F}$), and the robot must avoid cells identified as obstacles ($\mathcal{O}$). The robot is equipped with a costly sensor capable of identifying the actual status of unknown cells within a given radius, $r_s$, of its current location, $x_t$. This sensing region, $S(x_t)$, includes all cells $x$ that satisfy:

$$S(x_t) = \{x \mid \mathrm{dist}(x, x_t) \leq r_s\}, \tag{3.1}$$

where $\mathrm{dist}(\cdot, \cdot)$ is a distance metric applied within grid world $\mathcal{W}$. This thesis uses the Manhattan distance metric, resulting in a circle-shaped sensing region around $x_t$. The operation of the sensor is illustrated in Figure 3.1. When activated, the sensor reveals whether each cell within $S(x_t)$ is free or an obstacle. This information updates the robot's understanding of $\mathcal{W}$, allowing it to refine its path choices and adapt as new data is collected. The sensing decision at time $t$ is represented by $\theta_t \in \{0, 1\}$, where $\theta_t = 1$ indicates active sensor usage.

### 3.1.3 Information Structure

The robot relies on the knowledge it accumulates during navigation to inform its actions. At any time $t$, the robot's information about $\mathcal{W}$ is captured by the sets $\mathcal{F}_t, \mathcal{O}_t$, and $\mathcal{U}_t$, representing the cells known to be free, obstacles, or unknown, respectively.[1] These sets change over time as the robot gathers more information:

$$\mathcal{F}_t \supseteq \mathcal{F}_{t-1}, \quad \mathcal{O}_t \supseteq \mathcal{O}_{t-1}, \quad \mathcal{U}_t \subseteq \mathcal{U}_{t-1}. \tag{3.2}$$

This cumulative knowledge, denoted as $\mathcal{I}_t = \{\mathcal{F}_t, \mathcal{O}_t, \mathcal{U}_t\}$, provides the robot with an evolving understanding of its environment, shaping its navigation decisions. Given the robot's current location $x_t$, a planned path $\mathbf{p} = \{x_t, x_{t+1}, \ldots\}$ is deemed executable if every cell along $\mathbf{p}$ lies within the set of known free cells $\mathcal{F}_t$. Thus, the set of executable paths to the goal $g$ at time $t$, starting from $x_t$, is denoted as $\mathcal{P}(x_t, g, \mathcal{I}_t)$. This set of paths relies on the robot's sensing strategy and represents an under-approximation of all feasible paths to the goal. Paths within $\mathcal{P}(x_t, g, \mathcal{I}_t)$ are executable under the robot's current information structure, informing it of feasible routes given its limited knowledge at time $t$.

### 3.1.4 Traversal Cost

The traversal cost framework provides a flexible and adaptable means of defining the robot's path-planning objectives, allowing it to optimize for various criteria such as minimizing path length, avoiding detection, or adhering to mission-specific constraints. The cost of traversing a free cell $x \in \mathcal{F}$ from an adjacent cell is denoted as $c(x)$, while cells in $\mathcal{O}$ (obstacles) are considered untraversable, with their cost set to $c(\bar{x}) = \infty, \forall\, \bar{x} \in \mathcal{O}$. This cost formalism is intentionally kept vague, enabling it to be tailored to the specific requirements of different applications. For instance, $c(x)$ could represent the time required to traverse a cell, the risk of detection in a militaristic

[1] $\mathcal{W} = \mathcal{F}_t \cup \mathcal{O}_t \cup \mathcal{U}_t$ and $\mathcal{F}_t \cap \mathcal{U}_t = \mathcal{O}_t \cap \mathcal{U}_t = \mathcal{F}_t \cap \mathcal{O}_t = \emptyset$ for all $t$.

scenario, or a general measure of environmental difficulty.

The total cost of a given path $\mathbf{p} = \{x_1, \ldots, x_t, \ldots, x_H\}$ is calculated as the sum of the traversal costs for each cell along the path:

$$c(\mathbf{p}) \triangleq \sum_{t=1}^{H-1} c(x_{t+1}), \tag{3.3}$$

where $H$ represents the length of the path.

For simplicity in this thesis, $c(x)$ is defined as 1 for all $x \in \mathcal{F}$, reducing the framework to the objective of minimizing the length of the path. This simplification allows for straightforward testing and evaluation while demonstrating the methodology's core principles. However, the formulation is sufficiently general to accommodate a wide range of cost functions, enabling future extensions to more complex optimization objectives.

In addition to traversal costs, the total sensor activations along a path of length $H$ are given by:

$$\sum_{t=1}^{H} \theta_t, \tag{3.4}$$

where $\theta_t$ is a binary indicator of whether the sensor is activated at time step $t$. This metric quantifies the resource cost–such as energy or computation–incurred by reducing environmental uncertainty through sensing. Together, traversal and sensing costs provide a comprehensive framework for defining the robot's navigation objectives and balancing path efficiency with strategic sensor usage. This dual optimization highlights the versatility of the approach, making it applicable to various real-world scenarios with diverse constraints and goals.

## 3.2 Formal Problem Statement

With these definitions in place, the problem can be formally stated. Figure 3.3 shows an example of what this solution may look like for the environment in Figure 3.1.

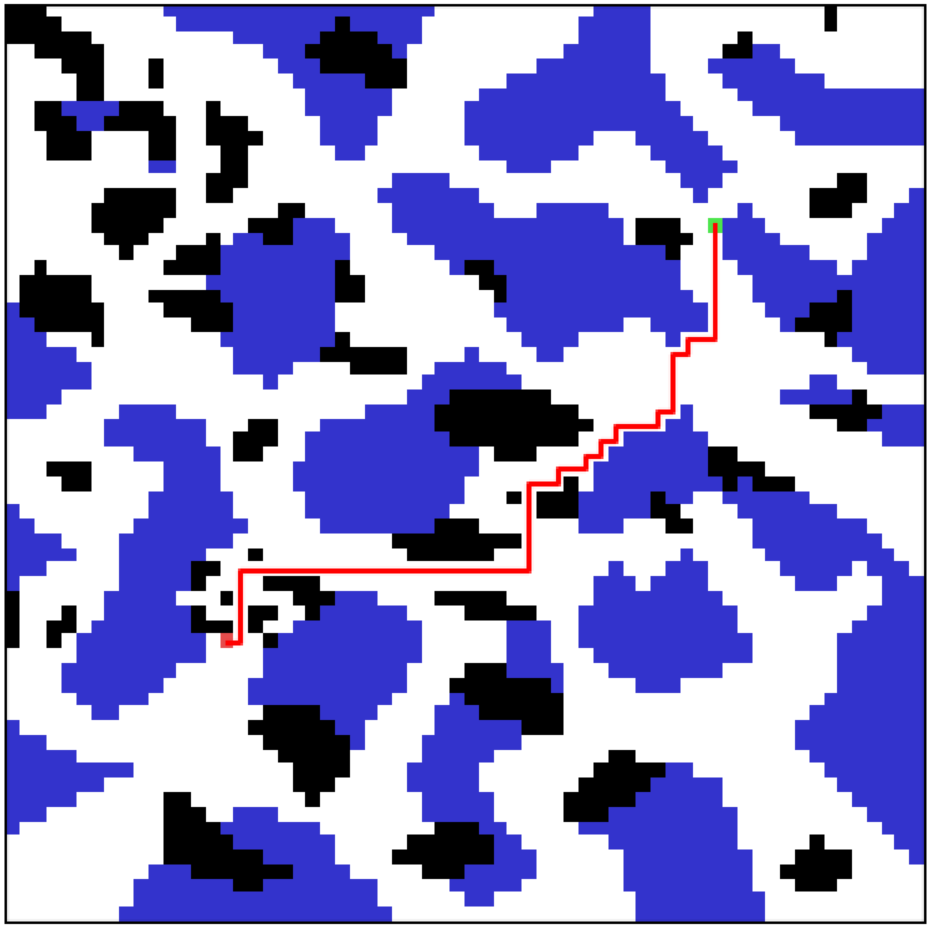


Figure 3.3: An example solution on the example world (See Figure 3.1). The start position is shown in the green square near the top right corner. The goal position is the red square near bottom left corner. The path taken is shown in red.

***Problem Statement:*** *For given start and goal locations, $s$ and $g$, respectively, and a sensing budget $N$, the objective is to find the optimal path from $s$ to $g$ while ensuring the robot remains within the free cells and uses the sensor no more than $N$ times.*

## 3.3 Assumptions

This is an incomplete information decision-making problem where the robot does not have prior knowledge about which locations are ideal for sensor activation. The robot also lacks knowledge of the obstacle density within the unknown regions, meaning its choices in both path and sensor usage directly impact the feasibility and cost of reaching the goal. Due to this incomplete information, finding a feasible path within the sensor budget is challenging. To ensure a well-posed problem, we introduce the following assumption:

**Assumption 1**: $\mathcal{P}(s, g, \mathcal{I}_0) \neq \emptyset$.

In other words, there exists at least one feasible path between the start and goal locations at $t = 0$. This assumption ensures the problem remains solvable within the given constraints, enabling the robot to navigate from the starting point to the goal while managing its limited sensing resources.

# CHAPTER 4: Preliminary Work

The preliminary work provides insights into the robot's decision-making process in two experimental setups: the World with a Single Unknown and Worlds with Scattered Unknowns. These scenarios illustrate the impact of varying levels of environmental complexity on the robot's strategy for managing uncertainty and sensor usage.

## 4.1 World with a Single Unknown Cell

The world with a single unknown cell i.e., $|\mathcal{U}_0| = 1$, is designed to explore the impact of a single point of uncertainty on the robot's path planning and sensor use. By focusing on a single unknown cell, this setup highlights the fundamental decision-making strategies—*pessimistic* and *optimistic*—and provides a simpler environment to analyze the effects of each strategy in isolation. Since only one cell is unknown at the start, we can simplify the problem without considering a sensing budget, allowing us to focus solely on the strategic implications of this uncertainty. This setup, illustrated in Figure 4.1, enables us to examine the robot's decision-making process when it navigates toward a goal with minimal unknown information.

### 4.1.1 Strategies: Pessimistic and Optimistic Paths

To address the unknown cell's uncertainty, the robot can adopt one of two fundamental strategies: a *pessimistic* path, which assumes that the unknown cell is an obstacle, or an *optimistic* path, which assumes the unknown cell is free. These strategies allow the robot to explore two opposing approaches to decision-making in uncertain environments.

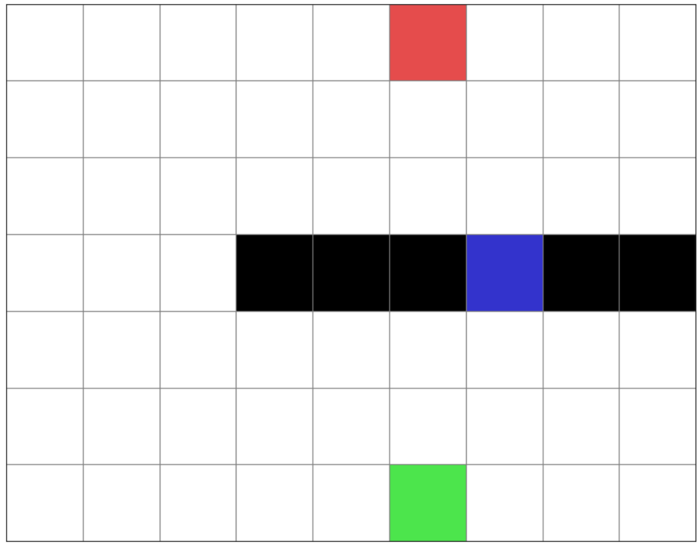


Figure 4.1: World with one unknown cell in it

**Definition 1 (Optimistic Path)** *An optimistic path is a path through* $\mathcal{W}$ *treating the unknown cells as free cells, allowing the robot to proceed as if all potential routes are traversable. While this approach can lead to shorter, more efficient routes if the assumptions prove correct, it risks encountering obstacles that would invalidate the planned path.*

The set $\mathcal{Q}(x, y, \mathcal{I}_t)$ represents all *optimistically executable paths* from a starting cell $x$ to a target cell $y$, given the current environmental knowledge of the robot $\mathcal{I}_t$. Unlike $\mathcal{P}(x, y, \mathcal{I}_t)$, which contains only feasible paths restricted to known free cells, $\mathcal{Q}(x, y, \mathcal{I}_t)$ extends this concept by including paths that traverse both known free cells ($\mathcal{F}_t$) and unknown cells ($\mathcal{U}_t$). This optimistic assumption treats unknown cells as traversable, enabling the robot to explore potential routes that could be feasible if the unknown regions turn out to be unobstructed.

Mathematically, $\mathcal{Q}(x, y, \mathcal{I}_t)$ is defined as the set of paths $\mathbf{q}$ such that:

$$\mathcal{Q}(x, y, \mathcal{I}_t) = \{\mathbf{q} \mid \mathbf{q} = \{x, x_1, \ldots, x_n, y\}, x_i \in \mathcal{F}_t \cup \mathcal{U}_t, \forall\, i \in [1, 2, \ldots, n]\}. \tag{4.1}$$

Here, each path $\mathbf{q}$ in $\mathcal{Q}$ consists of a sequence of cells starting at $x$ and ending at

$y$, where all intermediate cells $x_i$ belong to either the set of known free cells $\mathcal{F}_t$ or unknown cells $\mathcal{U}_t$. This formulation allows the robot to compute potential paths under optimistic assumptions and is the foundation for evaluating optimistic strategies in uncertain environments.

At any time $t$, the optimistic path to the goal $g$ from a cell $x \in \mathcal{F}_t \cup U_t$ is the path in $\mathcal{Q}(x, g, I_t)$ with the minimum cost:

$$\mathbf{p}_t^{\text{opti}}(x) = \text{argmin}_{\mathbf{p} \in \mathcal{Q}(x,g,\mathcal{I}_t)} c(\mathbf{p}). \tag{4.2}$$

The optimistic *cost-to-go* captures the entire cost from $x$ to $g$. This provides a quantitative cost for the cells. For a cell $x \in \mathcal{F}_t \cup \mathcal{U}_t$ is defined as:

$$c_t^{\text{opti}}(x) = \begin{cases} c(\mathbf{p}_t^{\text{opti}}(x)), & x \in \mathcal{F}_t \cup \mathcal{U}_t, \\ +\infty, & x \in \mathcal{O}_t, \end{cases} \tag{4.3}$$

Figure 4.2 presents the color-map of $c^{\text{opti}}$, representing the cost-to-go under the optimistic strategy. The gradient shows the shortest path cost from each cell to the goal, with darker colors (near black) indicating lower costs close to the goal and brighter colors representing higher costs. White cells signify regions that were not evaluated under the optimistic strategy, emphasizing areas beyond the robot's consideration in this strategy.

**Definition 2 (Pessimistic Path)** *The pessimistic path represents a cautious strategy in which the robot treats unknown cells as obstacles. This approach leads the robot to avoid unknown areas, potentially resulting in a longer but guaranteed safe path that avoids uncertainty entirely.*

At any given time t, the pessimistic path to the goal $g$ from a cell $x \in \mathcal{F}_t$ is the path

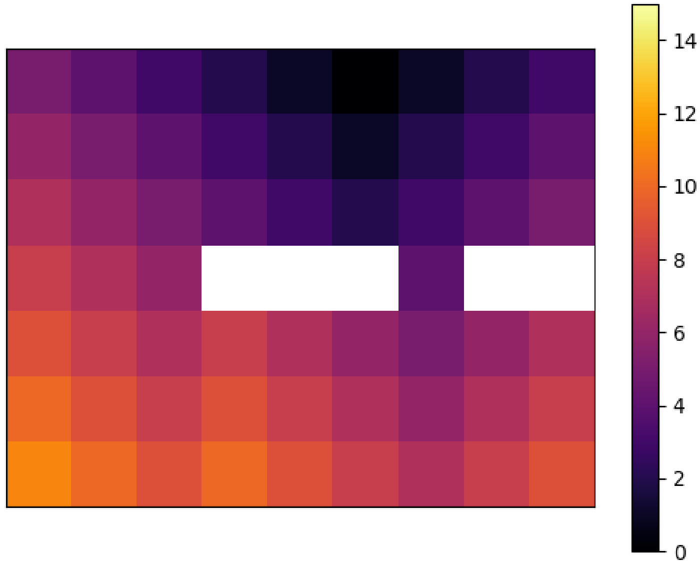



Figure 4.2: The color-map of $c^{\text{opti}}$.

in $\mathcal{P}(x, g, \mathcal{I}_t)$ with the minimum cost:

$$\mathbf{p}_t^{\text{pess}}(x) = \operatorname{argmin}_{\mathbf{p} \in \mathcal{P}(x,g,\mathcal{I}_t)} c(\mathbf{p}). \tag{4.4}$$

The pessimistic *cost-to-go* for a cell $x$ at time $t$ is defined as:

$$c_t^{\text{pess}}(x) = \begin{cases} c(\mathbf{p}_t^{\text{pess}}(x)), & x \in \mathcal{F}_t, \\ +\infty, & x \notin \mathcal{F}_t. \end{cases} \tag{4.5}$$

Figure 4.3 depicts the cost-to-go map $c^{\text{pess}}$ under the pessimistic strategy. The colors and meaning are identical to those in Figure 4.2 except this representation assumes all unknown cells are obstacles. This leads to longer paths and higher-value cells, reflecting the robot's avoidance of uncertain regions under this cautious strategy.

These two strategies represent contrasting risk management philosophies. The optimistic path tackles the unknown optimistically by assuming they are free and traversable cells, while the pessimistic path takes a conservative approach, assuming unknown cells contain obstacles. A distinct set of actions can be derived from $\mathbf{p}^{\text{pess}}$

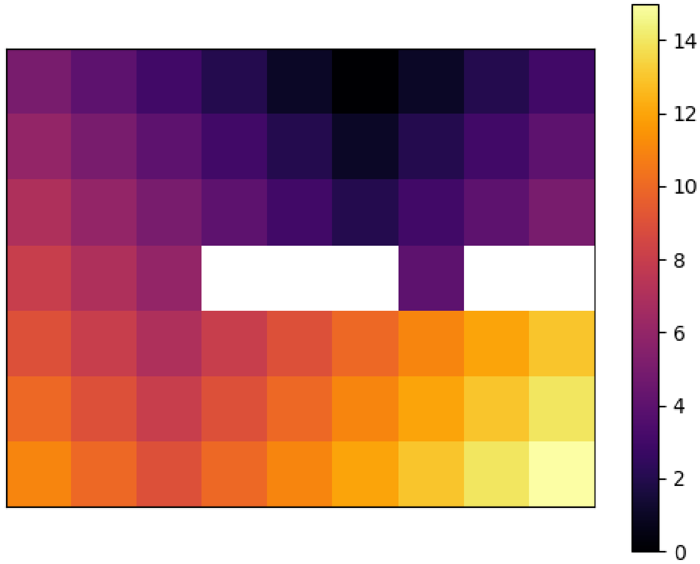
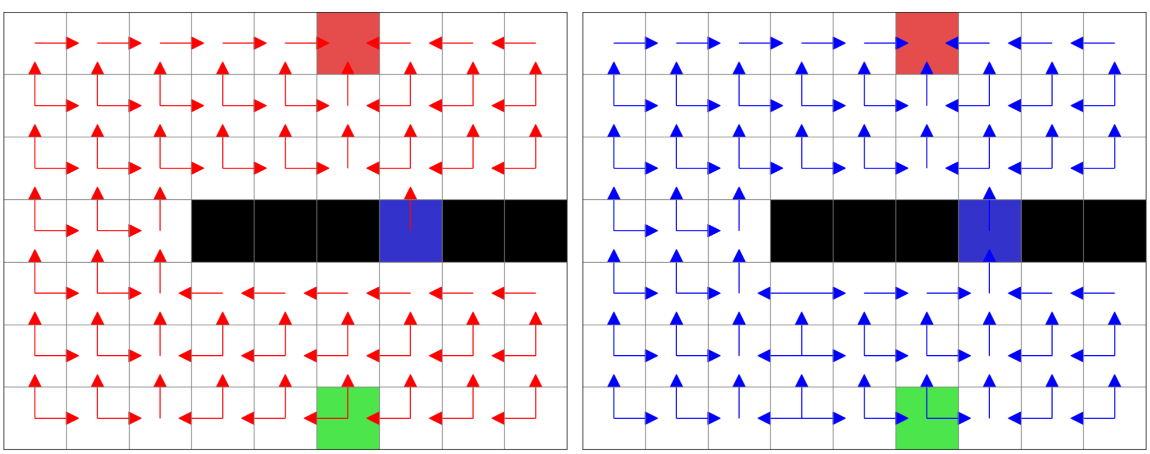



Figure 4.3: The color-map of $c^{\text{pess}}$.

and $\mathbf{p}^{\text{opti}}$, indicating the best actions to take at any cell. Thus, each strategy corresponds to a different *policy* or best actions to take when following either strategy. The optimistic approach, which assumes that all unknown cells are free, encourages the robot to take the shortest path, even through unknown regions, as shown in Figure 4.4b. Conversely, the pessimistic approach, shown in Figure 4.4a, assumes that unknown cells are obstacles and encourages the robot to avoid unknown areas entirely, taking a potentially longer but guaranteed safe route.



(a) Pessimistic policy.

(b) Optimistic policy.

Figure 4.4: Policy actions shown for both strategies.

### 4.1.2 Regret

In uncertain environments, *regret* provides a quantitative measure to evaluate the impact of decision-making on path efficiency [62]. It captures the trade-offs between optimistic and pessimistic strategies at a given moment and evolves as the robot gathers more information about the environment. Regret is particularly valuable in uncertain scenarios where the robot must balance risk and caution, as it reflects the potential benefit of the actions. Regret does not directly influence sensor usage or path decisions, but instead serves as a benchmark for understanding the consequences of past choices and the potential benefits of alternate strategies.

Formally, regret at a cell $x$ is defined as the difference between the pessimistic and optimistic path costs:

$$c_t^{\text{reg}}(x) = c_t^{\text{pess}}(x) - c_t^{\text{opti}}(x). \tag{4.6}$$

As the pessimistic path cost ($c_t^{\text{pess}}(x)$) assumes the path will only be within $\mathcal{F}_t$, and the optimistic path cost ($c_t^{\text{opti}}(x)$) assumes the path through ideal conditions through both $\mathcal{F}_t$ and $\mathcal{U}_t$, it follows that $c_t^{\text{pess}}(x) \geq c_t^{\text{opti}}(x)$. This relationship ensures that regret is always non-negative.

As shown in Figure 4.5, regret is calculated for every cell in the environment to reflect the difference in the costs of the path in each location. Positive regret indicates that the optimistic path would result in a shorter route to the goal compared to the pessimistic path. In other words, it quantifies the potential improvement in path length if the robot were to act optimistically from its current position onward. A high regret value may reflect suboptimality in the current decision-making process or highlight areas where sensing to follow an optimistic path could significantly improve efficiency.

Positive regret indicates that the optimistic path would result in a shorter route to the goal than the pessimistic path. In other words, it quantifies the potential


Figure 4.5: Color map of $c^{\mathrm{reg}}$ (see equation (4.6)).

improvement in path length if the robot were to act optimistically from its current position onward. A high regret value may reflect suboptimality in the current decision-making process or highlight areas where sensing to follow an optimistic path could significantly improve efficiency.

Zero regret occurs when the pessimistic and optimistic paths are identical in length, signaling that the robot has complete knowledge of the remaining environment to the goal. Further sensing does not affect the shortest available path, allowing the robot to proceed confidently without additional exploration.

Regret provides valuable insights into the robot's decision-making, but has limitations. It does not capture all aspects of the robot's strategy, such as how long-term goals or constraints like the sensing budget influence sensing decisions. Additionally, regret values alone cannot fully explain the interplay between exploration and efficiency, as they primarily reflect the difference between two predefined cost metrics. Despite these limitations, regret serves as an essential tool for evaluating the impact of uncertainty on path planning and identifying areas where improvements in decision-making could lead to better outcomes.

Through these interpretations, regret offers the robot a clear metric to guide decision-making. Positive regret highlights situations where exploration could be advantageous, and zero regret confirms that the robot's path is optimal under the current conditions. By tracking regret, the robot can dynamically adjust its strategy, balancing risk and efficiency to optimize its path toward the goal.

### 4.1.3 Simple Controller

The Simple Controller offers an efficient and streamlined approach to navigating the simple environment using the *intersection* of pessimistic and optimistic policies. By identifying actions where both policies agree, the controller provides the robot with a reliable path forward, minimizing the need for excessive sensor usage. This shared action set represents a valid best-case path regardless of the unknown cell's true nature.

As illustrated in Figure 4.6, the agreed-upon actions between the pessimistic and optimistic strategies are represented by green arrows. These green arrows indicate where both actions intersect, offering the robot a confident route to follow. By adhering to this path, the robot can proceed without activating its sensor, as the consistency between the strategies ensures a safe and reliable course of action.

However, situations arise where the two strategies disagree, leading to null sets of actions. We call these *indecision points*. These points, visually marked by yellow cells in Figure 4.6, occur when neither strategy offers a universally reliable action. At these locations, the robot faces uncertainty about how to proceed, as the lack of alignment between the optimistic and pessimistic strategies creates ambiguity.

To resolve this uncertainty, the Simple Controller activates the robot's sensor at these indecision points. By revealing the true status of the unknown cell, the sensor reduces environmental ambiguity and realigns the pessimistic and optimistic strategies. This realignment transforms the indecision point into an agreed-upon path, allowing the robot to continue along the unified route marked by green arrows.

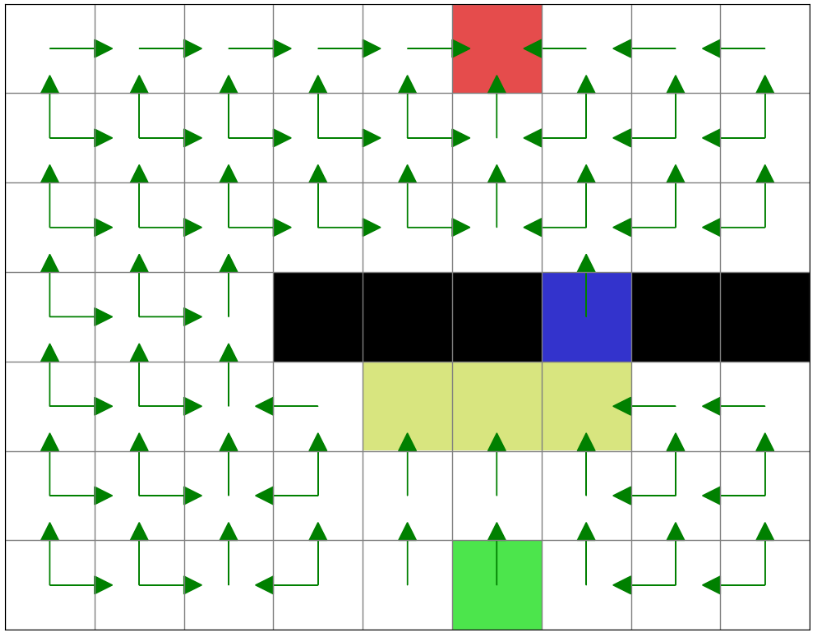


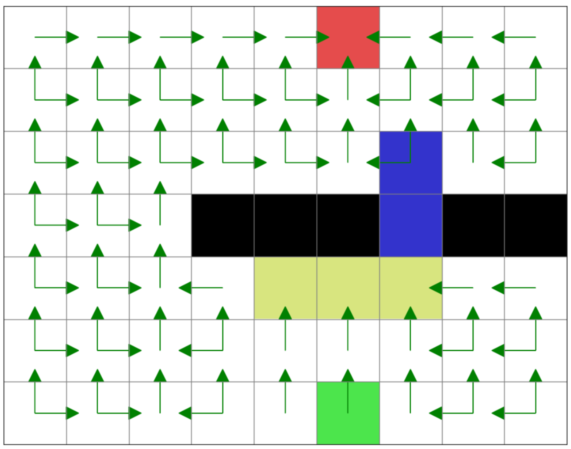
Figure 4.6: Agreed upon policy. See Figure 4.4 for Optimistic and Pessimistic Policies.

The Simple Controller thus achieves a balance between caution and efficiency. It follows the agreed path, bypassing sensor activation when the pessimistic and optimistic strategies intersect, and activates the sensor only when necessary at indecision points. By minimizing unnecessary sensor usage, the controller operates efficiently, making it particularly well-suited for environments with limited uncertainty, such as scenarios involving a single unknown. This approach ensures robust and effective control while optimizing the robot's resources and actions.

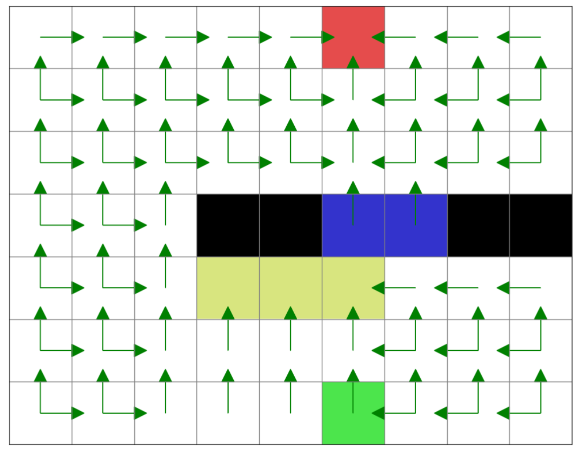
### 4.1.4 Single Group of Unknown cells

The scenario of a single group of unknown cells extends the principles explored in the case of a single unknown cell to environments where multiple adjacent cells are unknown. This setup introduces slightly more complexity by clustering uncertainty into a single contiguous region rather than isolating it to a single cell. Despite this increased complexity, the impact on the planning strategies remains identical, as the robot's fundamental decision-making process remains largely unchanged.

When navigating a single group of unknown cells, the robot still employs pessimistic and optimistic strategies to determine its path. The critical difference is changing the indecision points along the boundary of the unknown group. These indecision points arise as the robot evaluates whether to activate its sensor to reveal the group's status or to proceed cautiously along a longer, safer path. The increase in indecision points does not alter the underlying strategies but adds to the indecision points where sensor activation may be necessary to resolve ambiguity.



(a) Extending unknown vertically.



(b) Extending unknown horizontally.

Figure 4.7: Examples of the single unknown group.

From a decision-making perspective, the presence of a single group of unknown cells highlights the scalability of the approaches discussed for a single unknown cell. The Simple Controller, for instance, continues to function effectively by identifying intersections between pessimistic and optimistic strategies and activating the sensor only at critical indecision points. This ensures that the robot conserves its resources while making informed decisions that balance exploration and efficiency.

In summary, transitioning from a single unknown cell to a single group of unknown cells primarily introduces more indecision points, without fundamentally changing the robot's strategies or controller operation. This setup provides a stepping stone for understanding how the framework scales to more complex environments, reinforcing the robustness of the methods in managing clustered uncertainty while maintaining efficient navigation.

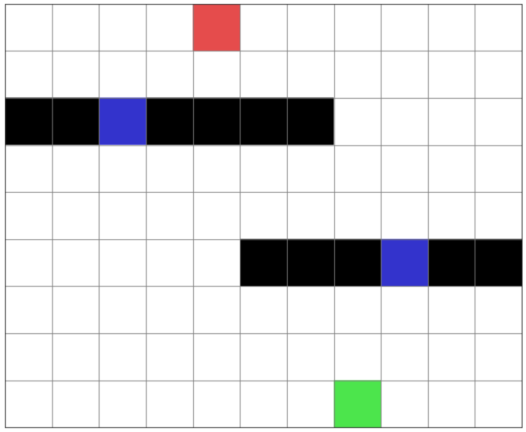
## 4.2 World with Scattered Unknown Cells

This section explores the challenges of navigating environments with scattered unknown cells, scaling up from simple cases to more complex scenarios. By introducing multiple areas of uncertainty, these experiments highlight the limitations of the simple controller and emphasize the need for strategies that account for sensor budget and global regret. As the number of unknown cells increases, the interconnectedness of regret and its global nature become apparent, along with the growing inefficiency of the simple controller in handling cumulative uncertainties.

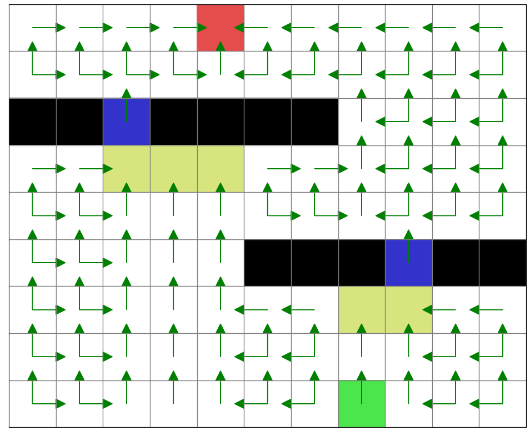
### 4.2.1 Scaling Up To Multiple Unknown Areas

Environments with two and three unknown cells introduce increasing complexity by adding layers of uncertainty and more indecision points. Figures 4.8 and 4.9 illustrate these setups, showing the worlds, their corresponding agreed-upon policies, and regret maps. For the two-cell environment (Figure 4.8), the robot navigates a zigzag obstacle pattern with two strategically placed unknown cells. The regret map (Figure 4.8e) shows distinct bands of regret forming around each unknown area, reflecting the potential inefficiencies caused by incorrect assumptions. These regret bands correlate directly with the agreed-upon policy (Figure 4.8d), where the simple controller consistently activates sensors at each indecision point to resolve uncertainty.

As the complexity increases with the three-cell environment (Figure 4.9), the challenges compound. The regret map (Figure 4.9c) demonstrates that regret is no longer isolated to individual unknown cells but becomes a global measure, capturing the cascading effects of decisions made at one location on other parts of the map. Each regret band aligns with a distinct policy action, emphasizing the interplay between local and global decision-making. The simple controller, while functional, continues to activate sensors at each indecision point without consideration of the overall sensing budget, resulting in inefficient sensor usage as the environment scales in complexity.



(c) The world.



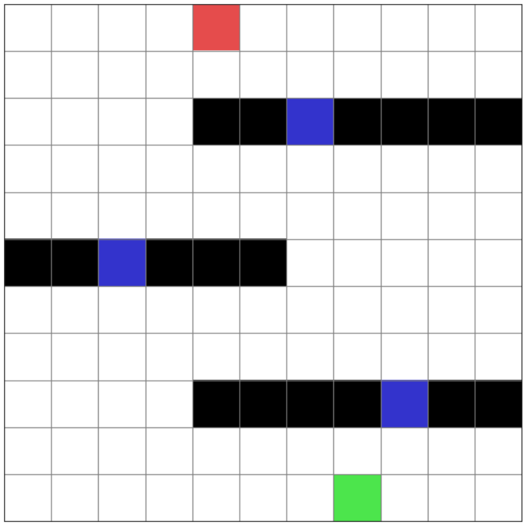
(d) Agreed upon policy.


(e) Regret color map.

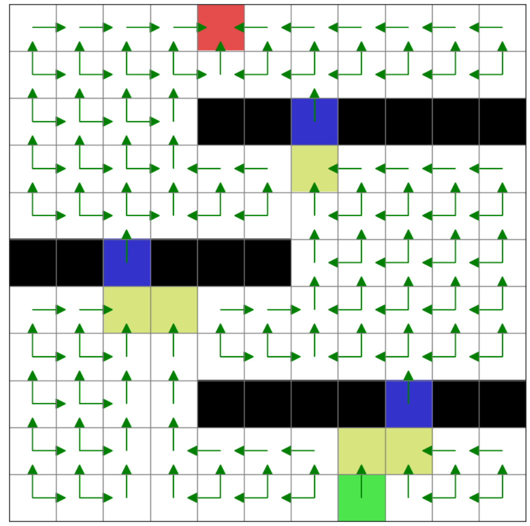
Figure 4.8: Environment with two unknown areas.

The regret maps across both scenarios emphasize regret's global nature. In the three-cell environment (Figure 4.9c), note how the regret near the bottom-right unknown cell mirrors the regret band above it. This visual indicates that decisions regarding one unknown area affect regret across the map, even in seemingly unrelated regions. These cumulative effects highlight the limitations of strategies that do not account for global implications when navigating environments with multiple unknowns.

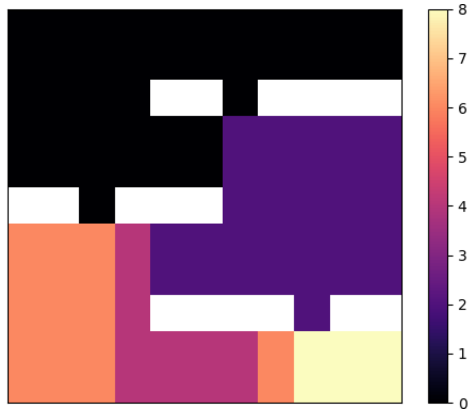
The simple controller, while effective in resolving local uncertainty at indecision points, lacks a mechanism to prioritize sensor activations strategically. It treats each indecision point independently, leading to suboptimal path efficiency and an inability



(a) The world.



(b) Agreed upon policy.



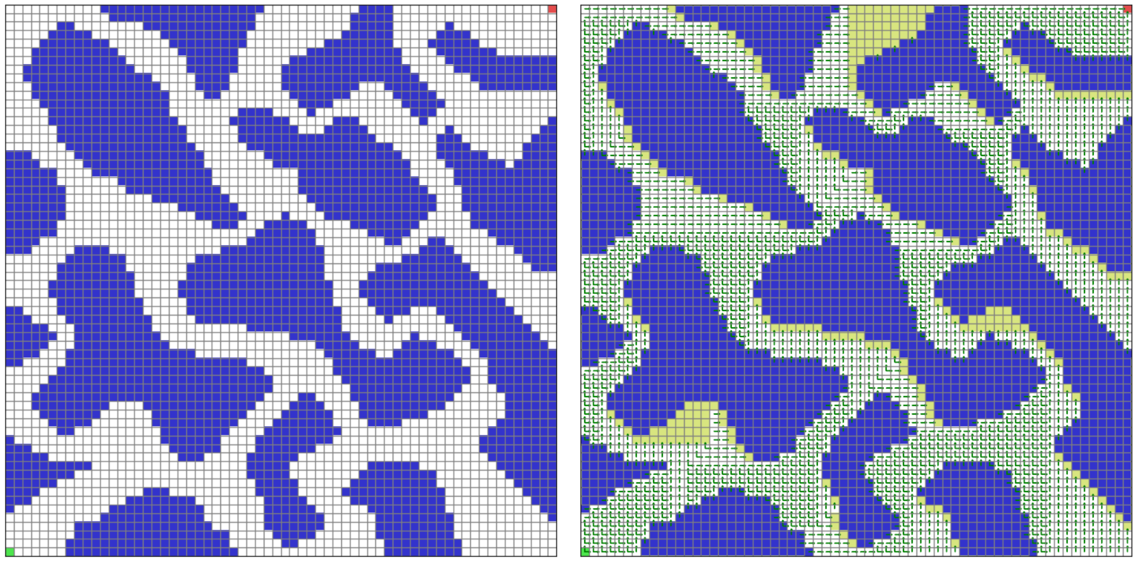

(c) Regret color map.

Figure 4.9: Environment for three unknown areas.

to conserve sensor budget. These findings motivate the need for advanced strategies that integrate global regret into decision-making, ensuring more effective navigation in environments with extensive unknown areas.

### 4.2.2 Larger Environment

In the larger, more realistic environment, illustrated in Figure 4.10a, is shown. The best case is shown for the starting portion of the map, illustrated in Figure 4.10b. The limitations of the simple controller become pronounced. This environment features extensive blotches of unknown regions with varying densities, presenting a highly



(a) The world. (b) Agreed upon policy.

Figure 4.10: Showcase of a more complex environment used for testing.

intricate decision-making landscape. The increased density of unknown cells results in a proliferation of indecision points, which overwhelms the simple controller. As some indecision points are positioned far from the robot's immediate path, activating the sensor at each one yields minimal or no new information, further straining the controller's effectiveness.

The regret map for this larger environment (see Figure 4.11) visualizes the compounded impact of uncertainty across numerous decision points, with high-regret regions highlighting the areas where incorrect assumptions could lead to costly detours or inefficient paths. In this scenario, the simple controller's strategy of treating each indecision point independently no longer suffices. The sheer number of unknowns creates an unsustainable demand for resources, and the robot's sensor activations lack prioritization, resulting in wasted energy and suboptimal navigation. The limitations of this approach reveal the need for a mechanism that can assess and prioritize indecision points based on their overall importance to the robot's mission.

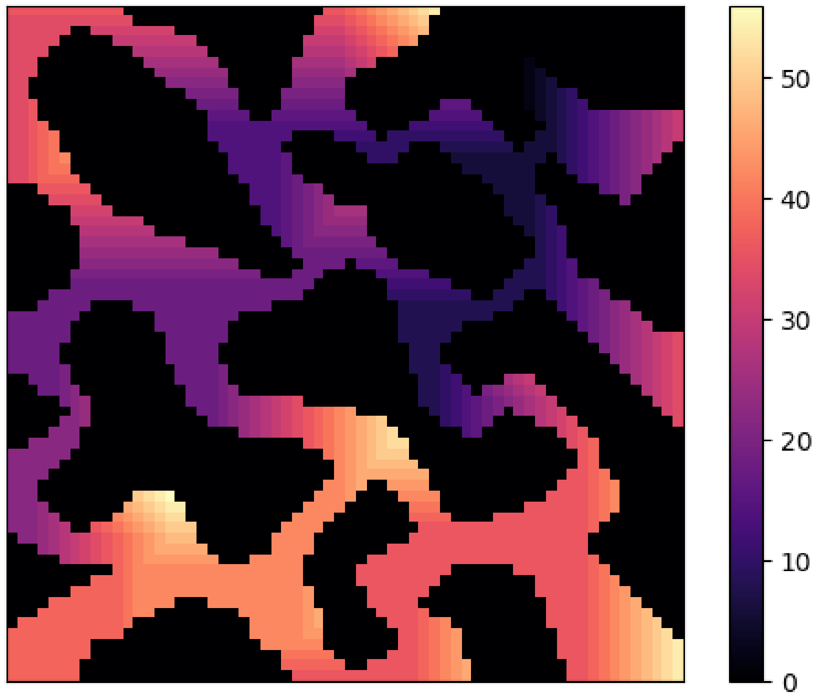



Figure 4.11: Regret colormap of the environment in Figure 4.10a.

### 4.2.3 The Need for Valuing Information: Toward a Utility/Cost Function

The scaled-up environments illustrate that the robot cannot afford to rely on a simple controller that activates sensors at every indecision point due to resource constraints and the increased number of decision points. As complexity grows, each sensor activation must be justified by its potential benefit to ensure efficient use of limited resources. This necessity reveals two primary factors for valuing information gained from sensor activations: information gain and path cost reduction.

First, information gain measures the number of unknown cells revealed by activating the sensor at a specific indecision point. In regions dense with unknown cells, a sensor activation that uncovers a significant portion of the environment offers high information value. By directing resources toward actions that yield substantial information gain, the robot can maximize the value of each activation, focusing on areas where the new information will most enhance its understanding of the environment. For instance, a sensor activation in a high-density area of unknowns could reveal key paths or obstacles that directly impact the robot's navigation choices, making this information especially valuable for path planning.

Second, path cost reduction assesses how sensor activation can lower the total cost of the robot's path by revealing shorter, feasible routes through unknown regions or by avoiding costly detours. This factor aligns with the goal of minimizing regret over areas with unknown information, as the regret map highlights regions where incorrect assumptions lead to inefficiencies in navigation. Sensor activations that reduce path cost by uncovering shorter, feasible routes help the robot avoid potential penalties associated with unverified assumptions about unknown cells. Therefore, prioritizing activations based on their potential to decrease path length supports the robot's overarching goal of achieving efficient and optimal navigation.

Together, these two factors-information gain and path cost reduction – underscore the necessity for a utility or cost function that can weigh each indecision point's relative importance to the robot's mission. Such a function would allow the robot to evaluate each activation choice according to both immediate and long-term benefits, enabling it to allocate sensor usage based on an assessment of total navigation efficiency rather than treating each indecision point in isolation. By valuing sensor activations according to these criteria, the utility/cost function would serve as a mechanism for balancing exploration and path optimization, supporting the robot's resource-limited operation in partially known environments.

## 4.3 Summary of Preliminary Work

The preliminary work demonstrates how the robot's decision-making process changes as environmental complexity increases. Beginning with a single unknown cell, the robot could experiment with simple strategies: a pessimistic approach that treats the unknown cell as an obstacle and an optimistic approach that assumes it is free. In this low-complexity setting, the simple controller proved effective, resolving uncertainty by activating the sensor at a single indecision point where the pessimistic and optimistic paths diverged. Here, the concept of regret-measuring the cost of incorrect assumptions – provided valuable feedback for evaluating the efficiency of each strategy, highlighting the costs associated with incorrect decisions.

As complexity increased with two and three unknown cells, more indecision points emerged, challenging the simple controller to manage multiple uncertainties. The regret map illustrated how each assumption's impact began accumulating across the environment, revealing regret as a global measure influencing overall navigation efficiency. The simple controller's binary sensor activation approach remained semi-functional in these moderately complex setups. However, it struggled to determine sensor usage with the large volume of indecision points.

The simple controller's limitations became evident in the larger, realistic environment with dense patches of unknown cells. The environment presented so many indecision points that activating the sensor at each one became inefficient, consuming resources without significantly improving navigation. The regret map for this environment highlighted regions where incorrect assumptions led to considerable cost penalties, underscoring the need for a more selective approach to sensor activation. The simple controller's inability to evaluate which indecision points held the highest value indicated that a more advanced prioritization mechanism was necessary.

These findings collectively reveal that, while the simple controller is effective in low-uncertainty environments, it cannot scale to handle dense or complex areas of unknown cells. This challenge points toward a utility or cost function to prioritize sensor usage. By assessing each potential activation based on information gain (number of unknown cells revealed) and path cost reduction (impact on total travel cost), a utility-based approach would enable the robot to make informed, strategic decisions about when to activate sensors. Such a framework could balance exploration with path efficiency, allowing the robot to manage uncertainty effectively and navigate complex environments with limited resources.

This summary of the preliminary work thus establishes the foundation for a more adaptive control strategy, paving the way for further research into utility-driven approaches that optimize information gathering and resource conservation.

# CHAPTER 5: Methodology 1: Regret-aware Joint Sensing and Path Planning

In this chapter, we present a decision-making framework that integrates path planning with sensor usage, allowing an agent to navigate through partially known environments while dynamically balancing exploration and efficiency. This joint sensing and control methodology leverages a real-time cost map that evaluates multiple metrics, guiding the robot's decisions on when and where to activate its sensor for maximum strategic value. The framework addresses limitations identified in the preliminary work by constructing a value-of-information (VoI)-based cost function, enabling the robot to bypass simple indecision points and actively pursue beneficial information.

## 5.1 Objective

The primary objective of this methodology is to develop an integrated framework for sensing and control, designed to navigate partially known environments effectively. Unlike traditional methods that address path planning and sensing independently, this approach seeks to combine these elements dynamically, balancing path length minimization with efficient sensor use. The ultimate aim is to address limitations observed in simpler controllers, especially in complex and partially mapped environments, where a naive approach to sensor usage may lead to inefficient or costly navigation.

In particular, this framework introduces a joint cost function that quantifies the value of information (VoI) gathered from each potential sensor activation. By assigning a calculable value to the information each cell might yield, the system determines optimal locations for sensor activation, thereby enhancing the robot's situational awareness and navigation efficiency. In essence, this joint sensing and control map is designed to adapt in real-time, making strategic use of the robot's limited sensing resources to reveal high-impact areas within the environment and navigate toward its goal with minimal unnecessary exploration.

## 5.2 Approach

This methodology introduces a joint framework for sensing and control that combines path planning and sensor activation within a unified system, dynamically adapting to real-time changes in the environment. By embedding a joint cost map, the framework equips the robot to make efficient, informed decisions on navigation and sensor usage, effectively addressing the complexities of partially unknown terrains.

Central to this approach is the *Unified Cost Map*, which integrates path planning with sensor activation strategies by quantifying the potential value of information (VoI) for each sensor activation. Each cell in this map has an associated cost, representing both real-time sensing information and anticipated travel costs. This cost map serves as a foundation for all subsequent path planning and exploration decisions, enabling the robot to prioritize paths and sensor activations that maximize overall efficiency and minimize redundant exploration.

The approach further emphasizes *Prioritization of Information Gain* to ensure that sensing actions are focused on regions of the map where the benefits of exploration are highest. By utilizing key metrics, including information gain, local regret, and estimated sensing requirements, the framework directs sensor resources toward cells that hold the most significant potential to improve path efficiency. This strategy reduces unnecessary exploration and aligns sensor usage with navigation objectives, enhancing the robot's ability to navigate effectively through unknown areas.

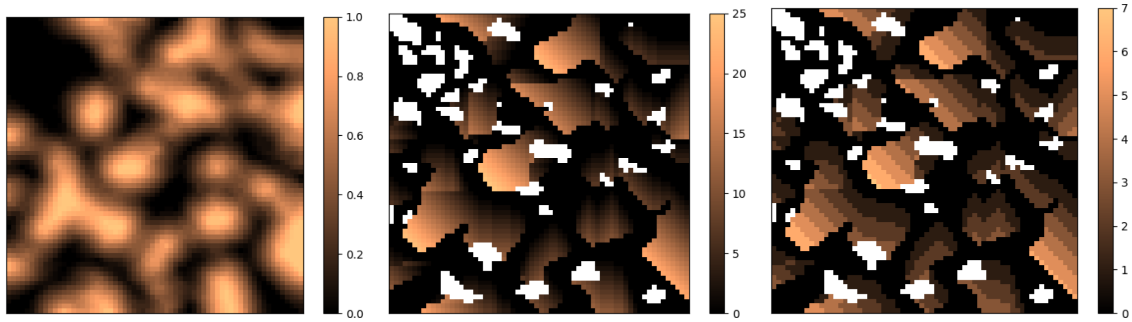
A critical feature of this framework is its *Weight-Driven Adaptability*, which allows the system to adjust to varying environmental conditions by tuning the influence of each metric on the cost map. The weights associated with each metric can be set to positive or negative values to either discourage or encourage specific actions. This tuning capability enables the system to meet diverse strategic goals, such as prioritizing exploration in highly uncertain areas or preserving resources by focusing on paths through known regions. Through this adaptability, the framework can optimize the balance between exploration and caution to achieve mission objectives more effectively.

As the robot progresses through its environment, the system performs *Dynamic Updates and Real-Time Sensing* to refine the navigation path based on newly acquired data. Each sensor activation feeds additional information into the cost map, allowing the robot to recalibrate its route based on the current understanding of the environment. When the sensor budget is exhausted, unknown cells are treated as effectively impassable, encouraging the robot to rely on known paths and avoid potentially risky unknown areas. This dynamic approach enables the robot to toggle between conservative and exploratory strategies as needed, optimizing both path length and sensor usage under evolving constraints.

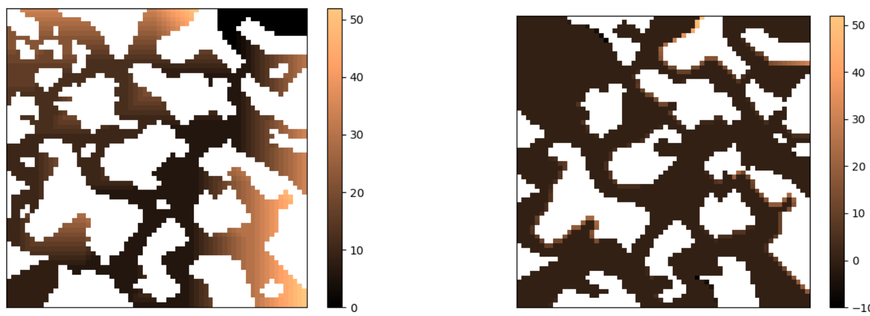
By combining adaptive cost mapping with real-time decision-making, this framework enables efficient exploration in complex environments, balancing path efficiency with information gain. Building on preliminary methods, this VoI-based approach addresses the limitations of simpler controllers, guiding the robot through partially known environments with enhanced precision, adaptability, and resource efficiency.

## 5.3 Information Metrics and Components

The joint sensing and control framework utilizes several key information metrics to inform the decision-making process as the robot navigates partially unknown environments. These metrics quantify the potential value of sensing, the trade-offs between optimistic and pessimistic path planning, and the costs associated with exploring or bypassing unknown regions. By integrating these components, the framework enables the robot to prioritize sensor usage and optimize its path through the environment




(a) $g^{\text{sense}}$ (b) $\ell$ (c) $c^{\text{est\_sense}}$




(d) $c^{\text{reg}}$ (e) $c^{\text{reg\_loc}}$

Figure 5.1: Graphs representing various cost components: (a) sensing gain (5.1), (b) estimated path length through unknown segment (5.2), (c) estimated sensing usage (5.5), (d) regret cost (4.6), and (e) local regret (5.4).

based on the expected information gain, regret costs, and potential sensing requirements.

### 5.3.1 Information Gain from Sensing

The information gain metric quantifies the amount of new data revealed when the robot activates its sensor. This is crucial in guiding the robot to strategically gather information in regions where unknown cells are likely to impact the path planning. Formally, the information gained at a cell $x$ at time $t$ is given by:

$$g_t^{\text{sense}}(x) = \frac{|S_t^u(x)|}{|S(x)|} = \frac{|S(x) \cap \mathcal{U}_{t-1}|}{|S(x)|}, \tag{5.1}$$

where $S(x)$ denotes the sensor's radius around cell $x$, $\mathcal{U}_{t-1}$ is the set of unknown cells at the previous time step, and $|\cdot|$ denotes the cardinality (or number of cells) in a set.

A value of $g_t^{\text{sense}}(x) = 1$ indicates maximum information gain, meaning that all cells within the sensor's radius were previously unknown. Conversely, a value of 0 implies that no additional information is obtained by sensing at that cell. Information gain, therefore, helps the robot prioritize sensor activations in regions where significant new data could be acquired, improving navigation decisions and potentially reducing detours.

### 5.3.2 Global Regret

Global regret measures the overall cost discrepancy between optimistic and pessimistic paths across the environment, providing insight into the cumulative impact of cautious versus exploratory decision-making. Regret[1] is a critical metric as introduced in Section 4.1.2 for quantifying the potential inefficiency introduced by uncertain regions, particularly in areas where unknown cells force the robot to choose between pessimistic and optimistic strategies. Global regret indicates the additional steps incurred by choosing a cautious approach (pessimistic strategy) over an adventurous one (optimistic strategy). High regret values suggest areas where the robot might benefit from further sensing, as the difference in path costs signals significant decision-making uncertainty. In essence, global regret measures the efficiency loss due to incomplete information and encourages exploration of regions with substantial potential cost savings.

### 5.3.3 Local Regret and Related Metrics

In contrast to global regret, *local regret* provides a focused measure of the uncertainty of decision-making for one region by not considering the global factor around particular cells. This metric extends the concept of regret from a global scope to a local context, allowing the robot to assess the impact of unknown regions directly surrounding it. We define several supporting metrics to calculate local regret, including

[1] Regret cost was defined as: $c_t^{\text{reg}}(x) = c_t^{\text{pess}}(x) - c_t^{\text{opti}}(x)$ in Eqn (4.6).

*unknown length* and *partial pessimistic cost*, which capture the traversal challenges and sensing requirements of areas near the robot's current position.

The first supporting metric, *unknown length*, denoted as $\ell_t(x_k)$, measures the distance across a contiguous segment of unknown cells that an optimistic path would need to cross to move from a known free cell to a goal. This value is calculated by following the optimistic path from a given cell $x$ back to the start location $s$, stopping at the first known cell $x_k$ encountered along the way. Formally, unknown length is expressed as

$$\ell_t(x_k) = |\mathbf{p}_t^{\text{opti}}(x \to x_k)|, \tag{5.2}$$

where $\mathbf{p}_t^{\text{opti}}(x \to x_k)$ represents the segment of the optimistic path from $x$ to $x_k$. The length $\ell_t(x_k)$ provides insight into the potential benefit of exploring unknown segments, as longer unknown paths suggest regions where sensor activation may significantly improve path efficiency. Figure 5.1b illustrates unknown length as a spatial distribution across the environment, helping the robot prioritize sensing in regions with extensive, contiguous unknown segments.

The *partial pessimistic cost*, $c_t^{\text{partial\_pess}}(x_k)$, combines the optimistic length through an unknown region with the pessimistic path to the goal, creating a hybrid metric that balances exploratory and cautious path planning. Defined as

$$c_t^{\text{partial\_pess}}(x) = \ell_t(x_k) + c_t^{\text{pess}}(x_k), \tag{5.3}$$

where $c_t^{\text{pess}}(x)$ is the pessimistic path cost to the goal, this metric enables the robot to explore unknown cells up to a certain point and then follow a safe, known path to reach the goal. By combining both optimistic and pessimistic approaches, partial pessimistic cost helps the robot make informed choices about when to activate sensors and when to avoid unknown areas. This balance allows the robot to approach navigation with both safety and efficiency, leveraging revealed information while managing the risks of unknown regions.

With these components in place, we define *local regret*, denoted as $c_t^{\text{reg\_loc}}(x_k)$, to quantify the potential benefit of sensor activation in the robot's immediate vicinity. Local regret is the difference between the full pessimistic path cost through $x_k$ and the partial pessimistic path through the same point:

$$c_t^{\text{reg\_loc}}(x) = c_t^{\text{pess}}(x) - c_t^{\text{partial\_pess}}(x). \tag{5.4}$$

A high value of local regret indicates that sensor activation at $x_k$ would likely reduce travel costs by revealing unknown cells that could shorten the overall path to the goal. This encourages sensor use in areas of high regret, where exploration would yield immediate path improvements. Figure 5.1e shows a map of local regret in the environment, illustrating potential locations where sensor activation could meaningfully improve navigation efficiency. By utilizing local regret and its supporting metrics, the robot can make targeted, data-driven decisions to explore specific unknown regions, ultimately balancing exploration with efficient path planning.

### 5.3.4 Estimated Sensing Requirement

The estimated sensing requirement metric provides an approximation of the number of steps required to traverse a given unknown area. Calculated as the ratio of the unknown length $\ell_t(x_k)$ to the sensor radius $r_s$, this metric prioritizes areas where a single sensor activation could reveal a large portion of unknown cells. Formally:

$$c_t^{\text{est\_sense}}(x_k) = \frac{\ell_t(x_k)}{r_s}, \tag{5.5}$$

This metric estimates the number of sensor steps necessary to explore an unknown region, helping the robot allocate sensor resources more effectively, particularly in complex environments with numerous unknown segments.

### 5.3.5 Transition Penalty

The transition penalty component discourages unnecessary entries into unknown regions unless there is a compelling reason to explore. Defined by the indicator function $c_t^{\text{tran}}(x)$, the transition penalty applies an additional cost to cells that lie on the boundary between known and unknown regions:

$$c_t^{\text{tran}}(x) = \begin{cases} 1, & \text{if } x \in \mathcal{U}_t \text{ and } \exists\, y \in \mathcal{F}_t \text{ s.t. } \operatorname{dist}(x, y) = 1, \\ 0, & \text{otherwise.} \end{cases} \tag{5.6}$$

This term is instrumental in shaping the robot's path to minimize unnecessary exploration, as it applies a cost for moving into the unknown unless the path planning objectives justify such an action. By incorporating transition penalty, the robot avoids high-risk areas without sufficient potential benefits from sensing, effectively balancing exploration and caution.

## 5.4 Mathematical Formulation of the Cost Map $C_t(x)$

The joint cost map $C_t(x)$ serves as the central framework for guiding the robot's navigation and sensing decisions, dynamically integrating various information metrics to balance exploration and path efficiency. Each cell $x$ in the environment is assigned a cost that reflects its classification (free, obstacle, or unknown) and the influence of specific metrics, such as information gain, regret, and estimated sensing requirements. The cost map is governed by the following piecewise function:

$$C_t(x) = \begin{cases} \infty, & \text{if } x \in \mathcal{O}_t, \\ w_0, & \text{if } x \in \mathcal{F}_t,\ c_t^{\text{pess}}(x) = c_t^{\text{opti}}(x), \\ w_1 + w_2 \cdot c_t^{\text{reg\_loc}}(x) + w_3 \cdot c_t^{\text{reg}}(x) + w_4 \cdot g_t^{\text{sense}}(x), & \text{if } x \in \mathcal{F}_t,\ c_t^{\text{pess}}(x) \neq c_t^{\text{opti}}(x), \\ w_5 + w_6 \cdot g_t^{\text{sense}}(x) + w_7 \cdot c_t^{\text{tran}}(x) + w_8 \cdot c_t^{\text{est\_sense}}(x), & \text{if } x \in \mathcal{U}_t,\ N_t < N, \\ \infty, & \text{if } x \in \mathcal{U}_t,\ N_t = N. \end{cases} \tag{5.7}$$

Here, the weights $w_0$ through $w_8$ act as tuning parameters that control the influence of each term in the cost function. While their specific roles and effects are thoroughly discussed in the subsequent chapter, they are introduced here as they appear explicitly in $C_t(x)$.

The cost map $C_t(x)$ evolves as the robot explores its environment, with costs updating dynamically based on newly acquired information. This adaptability enables the robot to adjust its navigation strategy in real time, reflecting changes in both known terrain and strategic priorities. For example, as sensor activations reveal unknown areas, the robot's path planner recalculates the cost map, balancing the immediate cost of exploration against the long-term benefits of revealing new information.

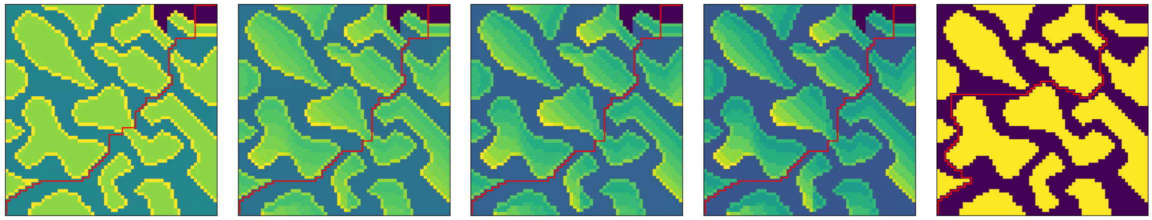
## 5.5 Example of $C_t(x)$ Adaptation

Figure 5.2 illustrates the dynamic behavior of $C_t(x)$ in a test environment as the number of sensors used ($N_t$) changes. The series of subfigures demonstrates how the cost map evolves over time, with the weights influencing the planner's decision-making process. The red line represents the planned path, which adapts to the updated costs as more sensors are used. Importantly, while the map shapes remain constant due to unchanged environmental knowledge ($I_t$), the cost values differ, reflecting the planner's dynamic prioritization based on sensor usage.

Notably, in Figure 5.2e, the yellow cells are assigned infinite cost ($C_t(x_t) = \infty$ where $x_t \in \mathcal{F}_t \bigcup \mathcal{U}_t$) once the sensing budget is exhausted ($N_t = N$). At this point, the planner defaults to basic A* behavior, treating all remaining unknown cells as obstacles. This ensures that the robot prioritizes safe navigation when additional sensing is no longer feasible.

## 5.6 Explanation of Weights and Their Roles

The weights $w_0$ through $w_8$ play a critical role in shaping the behavior of the cost map by defining the impact of various metrics on the robot's decision-making process.



(a) $N_t = 0$ (b) $N_t = 1$ (c) $N_t = 2$ (d) $N_t = 3$ (e) $N_t = 4$

Figure 5.2: Cost map of the environment, with the proposed path marked by the red line. The color scale represents the cost values, where purple indicates the lowest cost regions and yellow indicates the highest cost regions. Note that the path doesn't change between (b) and (d) because new info has been revealed.

These weights can be positive or negative, which affects how the robot interprets costs and incentives within the environment.

Positive weights act as cost additions, effectively discouraging the robot from selecting actions or paths associated with the respective metric. For example, a positive weight on global regret (e.g., $w_3$) increases the cost in areas of high regret, thereby discouraging routes where the pessimistic and optimistic paths diverge significantly. This positive weight drives the robot toward more conservative decisions, favoring known, safer paths. Likewise, a positive transition penalty weight (e.g., $w_7$) raises the cost of transitioning into unknown areas, signaling the robot to avoid entering such regions unless absolutely necessary. Positive weights can thus enforce a cautious approach, ensuring the robot remains in safer zones while using sensors sparingly and only where they add significant value.

Conversely, negative weights act as cost reductions, incentivizing exploration and the use of sensors in areas with promising metrics. Negative weights provide a 'discount' on cost in regions where the corresponding weighted metric is high, promoting the exploration of unknown regions when the potential gains justify the use of the sensor. For instance, a negative weight on information gain (e.g., $w_4$) lowers the cost for cells with high information potential, encouraging sensor activations in these regions. Similarly, a negative weight on local regret (e.g., $w_2$) decreases the cost in areas of high local regret, where bypassing unknown cells could lead to missed path improvements. By reducing the costs associated with these areas, negative weights enable the robot to take calculated risks and prioritize information gathering, especially when the likelihood of improved path efficiency or reduced overall regret outweighs the potential downsides of exploration.

The weights can also be dynamically adjusted over time to enhance adaptability. For example, weights could be functions of the sensing budget $\frac{N_t}{N}$, where $N_t$ is the number of sensors used at time $t$ and $N$ is the total sensing budget. Which would allow the robot to prioritize certain actions as its sensing resources are depleted. This dynamic weight adjustment could be beneficial when the robot needs to be increasingly cautious or conservative as it approaches its sensing budget limit.

Each weight plays a distinct role in guiding the robot's navigation and decision-making in the joint sensing and control map. The weight $w_0$ represents the base traversal cost for regions with no potential for path improvement, specifically cells with identical optimistic and pessimistic paths. These cells are fully known, making them safe for traversal without requiring further exploration or sensor activation. The cost associated with $w_0$ maintains consistency across known areas, offering a stable base for path planning.

The weight $w_1$ applies a base traversal cost to regions with potential path improvement. In these cells, the optimistic and pessimistic paths diverge, indicating that additional exploration might reveal a more efficient route. As the base cost for these areas, $w_1$ signals the possibility of discovering shorter paths, while other factors, such as local regret and information gain, may further influence the decision to explore. This setup allows the robot to evaluate whether the potential for path improvement justifies sensor activation in regions of unknown cells.

The local regret weight $w_2$ adjusts the impact of the local regret metric $c_t^{\text{reg-loc}}(x)$, which quantifies missed opportunities associated with not exploring nearby unknown cells. When this weight is set higher, it encourages the robot to prioritize areas with substantial local regret, where sensor usage could prevent costly detours and optimize the path. By promoting exploration where it is likely to yield better routes, $w_2$ plays a crucial role in identifying high-value areas for sensor activation, focusing on local regions that offer immediate path efficiency gains.

The global regret weight $w_3$ modulates the influence of global regret $c_t^{\text{reg}}(x)$, representing the accumulated penalty for paths that traverse high-regret zones. These zones have significant discrepancies between the optimistic and pessimistic paths, suggesting a high degree of uncertainty. Positive values of $w_3$ discourage the robot from taking paths through these high-regret areas, thereby enforcing a more conservative strategy where long-term navigation safety is prioritized over potentially risky shortcuts. By assigning cost penalties to high-regret regions, $w_3$ supports the robot's cautious approach to areas with considerable unknown elements.

Information gain weights, specifically $w_4$ for free cells and $w_6$ for unknown cells, encourage sensor activation in areas with high information potential. These weights control the value assigned to newly revealed cells when sensors are activated. If these weights are negative, they reduce the cost for cells with significant information gain, guiding the robot to prioritize exploration in areas that yield substantial insights into the environment. By distinguishing between free and unknown cells, $w_4$ and $w_6$ adapt sensor usage to the type of region, ensuring sensors are deployed strategically where new information is most valuable.

The base cost for unknown cells with sensor access is set by $w_5$. This weight reflects the default cost of entering unknown areas and can be adjusted to either encourage or discourage exploration. A lower value for $w_5$ promotes exploration by lowering the barrier to accessing unknown regions, while a higher value limits sensor use to only essential situations. By setting an appropriate baseline cost for unknown areas, $w_5$ helps balance exploration with caution, ensuring sensors are used judiciously based on the surrounding context.

The transition penalty weight $w_7$ applies cost adjustments to transitions into unknown regions, especially near the boundaries of known cells. This weight discourages unnecessary entries into unknown regions, particularly where the potential for valuable information is low. By adding cost to boundary cells adjacent to known regions, $w_7$ moderates the robot's movements, reducing the likelihood of arbitrary sensor activations near familiar areas. This weight is particularly effective for managing boundary control between known and unknown regions, conserving sensor usage for areas with higher information value.

Lastly, the estimated sensing length weight $w_8$ assigns cost to the estimated sensing length $c_t^{\mathrm{est-sense}}(x)$, a metric representing the potential benefit of revealing longer unknown segments. By prioritizing paths where a single sensor activation can reveal larger unknown regions, $w_8$ ensures efficient use of the robot's sensor budget. This weight allows the robot to weigh the impact of each sensor activation, selecting cells where revealing a continuous unknown segment provides meaningful benefits for path planning and situational awareness.

Altogether, these weights form a balanced system within the joint sensing and control map, enabling the robot to navigate complex environments by integrating cautious planning with strategic exploration. Each weight's configuration can be tailored to specific mission goals, ensuring that the robot's path choices are well-informed and aligned with environmental constraints and resource availability.

## 5.7 Controller Operation: Flow and Execution

This section outlines the specific steps and strategies that guide the robot's navigation, from initialization through to reaching the goal. The controller operates by leveraging the joint cost map, which dynamically combines path planning with sensor activation to navigate through environments with partial knowledge.

To begin, the initialization process sets the framework for the robot's operation, defining the starting point $s$, goal $g$, and a sensing budget $N$ that limits the total number of sensor activations available. The joint cost map $C_t(x)$ is initialized with weights $w_0$ through $w_8$, each representing a different aspect of the environment's cost structure. These weights are input parameters that can be fine-tuned for specific conditions or strategic goals, such as minimizing path length or maximizing information gain. The initial path is generated using the A* algorithm (details of A* operation are outlined in Appendix A), which finds the least-cost path from $s$ to $g$ based on the initial cost map configuration. This path is then stored in a queue, allowing the robot to begin its journey towards the goal.

Each sensor activation triggers cost map updates as the robot proceeds along this path. When the sensor is activated, it reveals information about nearby unknown cells, which is used to update the cost map dynamically. By recalculating the values for affected cells, the controller can adapt to newly revealed obstacles or free paths, recalibrating the planned route to prioritize areas with lower costs or higher information value. If the sensing budget $N$ is exhausted, the cost of unknown cells is updated to infinity (see Equation (5.7)), effectively making them untraversable obstacles. This fallback strategy ensures that, in the absence of sensing, the robot defaults to a conservative planning approach where unknown cells are treated as obstacles.

During its journey, the robot continuously evaluates each next move in the path, checking whether the upcoming cell lies within a high-cost region influenced by unknown cells, global or local regret, or estimated sensing needs. When transitioning into an unknown area with a high associated cost, the robot evaluates whether to activate the sensor based on a predefined cost threshold. If the sensor is activated, the controller updates critical metrics based on the new data, including local and global regret, information gain, and estimated sensing requirement. After updating these values, the cost map $C_t(x)$ is recalculated, and the robot re-plans its path from the current position. This iterative recalibration ensures the robot makes informed choices in real time, balancing exploration with path efficiency.

The controller's operations are demonstrated in a series of figures that show the evolving cost map and the robot's path as it progresses through the environment. These figures highlight how the joint sensing and control map adapts with each sensor activation, illustrating the recalculated path and areas where sensor usage has influenced the decision-making process. Additionally, snapshots of the cost map at various stages of the journey show the changing cost structure as new information is gathered, visually demonstrating the cost adjustments that guide the robot's adaptive navigation.

The system's performance is evaluated using three key metrics: the pessimistic ratio, the optimistic ratio, and the closeness-to-optimal ratio. The pessimistic ratio is calculated as the actual path length divided by the pessimistic path cost $c_0^{\text{pess}}(s)$, which measures the system's effectiveness relative to an entirely cautious approach. A lower pessimistic ratio indicates that the controller makes intelligent path selections by leveraging sensing to avoid unnecessary detours. The optimistic ratio is the actual path length divided by the optimistic path cost $c_0^{\text{opti}}(s)$. The optimistic ratio provides insight into the efficiency of path choices; it is less reliable in heavily unknown environments due to the optimistic assumption of free unknown cells. Finally, the closeness-to-optimal ratio compares the controller's decisions relative to both the pessimistic and optimistic paths, offering a comprehensive view of the balance between exploration and path efficiency.

In summary, the controller operation is structured around an adaptive joint cost map, which combines sensing and path planning to guide the robot through uncertain environments. By dynamically recalculating path costs based on real-time information, the controller optimizes the use of sensors and navigational choices, achieving a balanced approach to exploration and efficiency. The adaptive nature of the system, coupled with the use of performance metrics, demonstrates the controller's ability to make intelligent and resource-conscious decisions in complex, partially known environments.

## 5.8 Performance Evaluation Metrics

The effectiveness of the joint sensing and control system is quantitatively assessed using several key performance metrics. These metrics evaluate the system's path efficiency, sensor usage, and adaptability in navigating partially known environments. By comparing the actual path with theoretically optimal and conservative paths, these metrics provide a comprehensive view of the controller's ability to balance exploration and efficiency.

The *pessimistic ratio* measures the actual path length relative to a conservative path that assumes all unknown cells are obstacles. Mathematically, this ratio is defined as:

$$\text{Pessimistic Ratio} = \frac{H}{c_0^{\text{pess}}(s)}, \tag{5.8}$$

where $c_0^{\text{pess}}(s)$ represents the pessimistic path cost from the starting location $s$ to the goal and $H$ denotes the actual path length, this ratio offers a baseline for evaluating the controller's effectiveness in circumventing cautious routes by leveraging sensor information.

A lower pessimistic ratio indicates efficient decision-making, allowing the controller to reduce path length by making selective sensor activations rather than avoiding unknown cells entirely. This metric is particularly useful in environments with a large amount of unknowns, where the system's ability to navigate through unknown regions drastically reducing the path length is crucial.

The *optimistic ratio* serves as a comparison between the actual path length and a theoretical path that treats all unknown cells as free. This optimistic path cost is represented as $c_0^{\text{opti}}(s)$, where the ratio is:

$$\text{Optimistic Ratio} = \frac{H}{c_0^{\text{opti}}(s)}, \tag{5.9}$$

where $c_0^{\text{opti}}(s)$ represents the optimistic path cost from the starting location $s$ to the goal $g$ and $H$ denotes the actual path length. This ratio measures the efficiency relative to an idealized, risk-tolerant strategy. While this metric captures the theoretical best-case scenario, it may overestimate efficiency in dynamic environments, as the optimistic path is not always feasible due to potential obstacles within unknown regions. However, it is valuable for comparing the system's performance against an idealized benchmark, particularly in environments where exploration costs are low and risk can be tolerated.

The *closeness-to-optimal ratio* provides a comprehensive metric for assessing the controller's ability to balance caution and exploration. This ratio evaluates how much closer the actual path is to the optimistic strategy than the pessimistic one, offering insight into the system's effectiveness in navigating uncertainty. Ideally, a value near 0% indicates that the chosen strategy aligns more closely with the optimistic path, demonstrating efficient exploration and reduced conservatism. In contrast, values closer to or exceeding 100% suggest a reliance on the more cautious, pessimistic approach.

This metric is particularly sensitive to system setting variations, such as the sensing budget and the true nature of unknown cells. Significant changes in these factors can dramatically alter the ratio, making it a valuable but context-dependent measure. Although not an absolute determinant, the closeness-to-optimal ratio serves as a useful indicator of the qualitative success of the method in striking the right balance

Table 5.1: Weight values for the cost function

| variable | $w_0$ | $w_1$ | $w_2$ | $w_3$ | $w_4$ | $w_5$ | $w_6$ | $w_7$ | $w_8$ |
|---|---|---|---|---|---|---|---|---|---|
| value | 0.2 | 1 | $-0.3$ | 1.5 | $-0.1$ | 1.75 | $-1$ | 0.3 | $2.5 \cdot N_t$ |

between optimistic and pessimistic strategies. The ratio is formally defined as:

$$\text{Closeness-to-optimal Ratio} = \frac{H - c_0^{\text{opti}}(s)}{c_0^{\text{pess}}(s) - c_0^{\text{opti}}(s)} \tag{5.10}$$

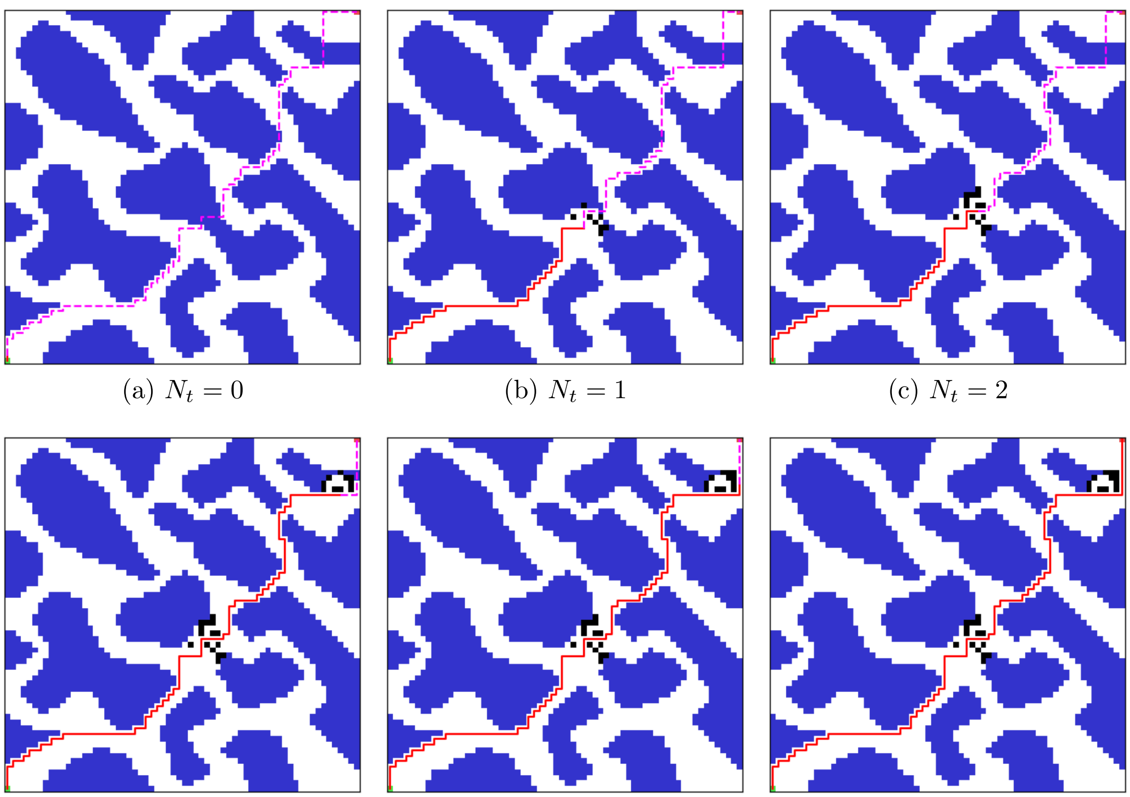
where $H$ is the actual path length, $c_0^{\text{opti}}(s)$ represents the optimistic path cost, and $c_0^{\text{pess}}(s)$ represents the pessimistic path cost. A ratio near zero percent signifies a strategy closer to the optimistic approach. In contrast, values approaching or greater than 100% reflect a more cautious, pessimistic strategy or the need to back down due to the information revealed.

Together, these metrics provide a robust framework for assessing the controller's navigation and sensing decisions in complex, uncertain environments. Each metric offers unique insights into path efficiency, sensor use, and adaptability, comprehensively evaluating the controller's performance across different operational scenarios. These quantitative metrics enable the fine-tuning of the joint cost map, optimizing path selection while balancing exploration with resource conservation.

## 5.9 Results

This study tested the joint sensing and control framework using a series of environments to evaluate its capacity to balance path efficiency with intelligent sensor usage in uncertain settings. The framework was tuned using a set of calibrated weights, which were essential to directing the agent's behavior by controlling the relative importance of each component in the cost map. These final weights, listed in Table 5.1, were adjusted to balance exploration and caution, with positive weights discouraging risky paths and negative weights encouraging sensor activation in information-rich regions.

The initial tests were conducted in a controlled map environment (see Figure 5.5a), which provided a baseline for observing the agent's behavior (see Figure 5.3) and the dynamic adjustments of the cost map (see Figure 5.4). In this setting, the agent's cost map was updated in real-time with each sensor activation, highlighting the adaptive nature of the framework. A sequence of figures illustrates how the cost map evolves with each new activation, incorporating newly revealed information and recalculating the path accordingly. This sequence demonstrates the framework's ability to prioritize sensor use at strategic locations, optimizing both path efficiency and resource usage. Through this iterative updating process, the framework enabled the agent to navigate efficiently by responding to new environmental insights while minimizing detours.

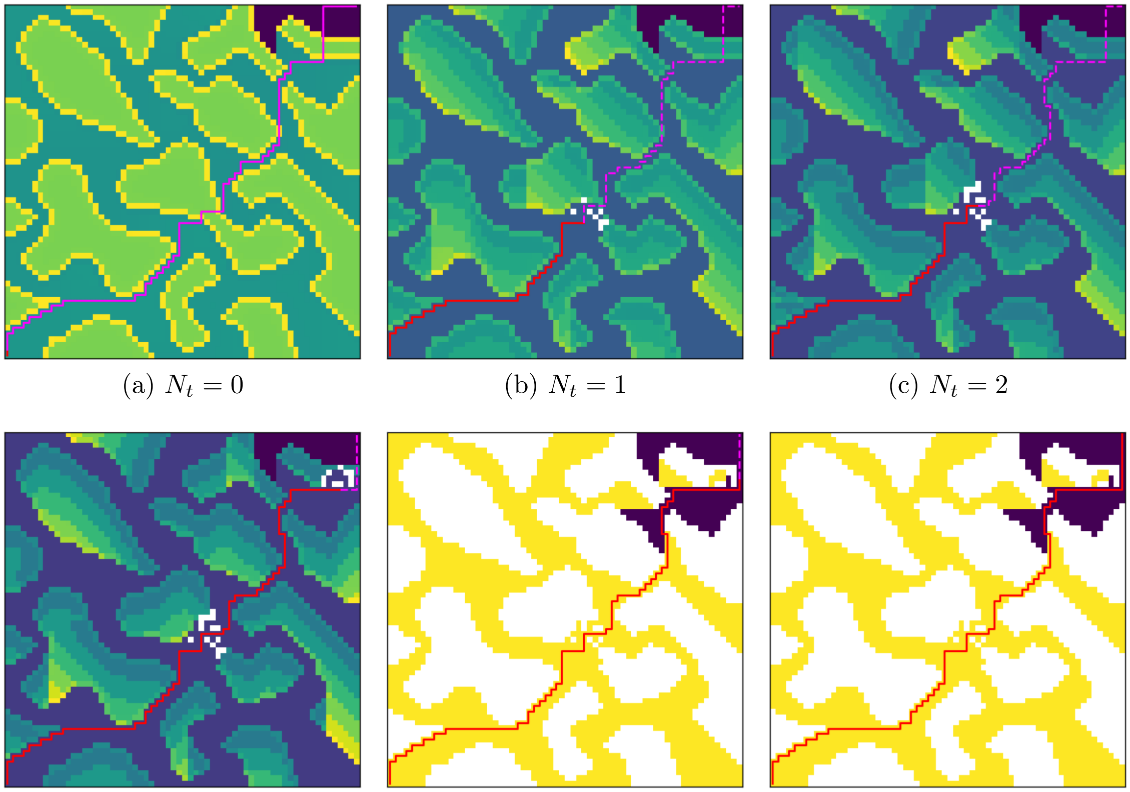


(a) $N_t = 0$ (b) $N_t = 1$ (c) $N_t = 2$

(d) $N_t = 3$ (e) $N_t = 4$ (f) final path

Figure 5.3: Demo of the as the controller ($N = 4$, $s =$ bottom left in green, $g =$ bottom right in right) navigates the environment with snapshots after every sensor activation. The path is in red, while the proposed path is the purple dashed line.

Following these initial tests, the framework was applied to four more complex eval-



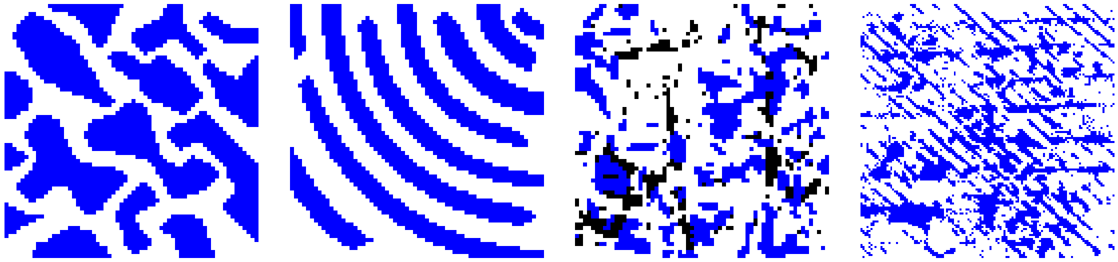
(a) $N_t = 0$ (b) $N_t = 1$ (c) $N_t = 2$

(d) $N_t = 3$ (e) $N_t = 4$ (f) final path

Figure 5.4: Demo of the cost map as the controller ($N = 4$, $s =$ bottom left in green, $g =$ bottom right in right) navigates the environment with snapshots after every sensor activation. The path is in red, while the proposed path is the purple dashed line.

uation environments, illustrated in Figure 5.5. Each environment presented unique challenges, from dense clusters of obstacles to scattered unknown areas that required the agent to balance exploration with caution. The diverse layouts in these maps allowed for comprehensive testing of the framework's adaptability, especially under conditions where sensor usage had to be rationed carefully to avoid depletion. In each of these environments, the framework was tasked with finding the most efficient path while managing limited sensor activations, simulating realistic constraints that could be encountered in exploratory missions with finite resources.

To provide a robust measure of the framework's performance, 500 trials were conducted in each of the evaluation environments, with the results averaged to yield reliable performance metrics. These trials measured three primary ratios: the *pes-*



(a) Env 0 (b) Env 1 (c) Env 2 (d) Env 3

Figure 5.5: Environments used to evaluate the proposed framework.

Table 5.2: Results averaging the pessimistic ratio, optimistic ratio, and closeness-to-optimistic ratio for methodology 2

| | Optimistic Ratio | Pessimistic Ratio | Closeness-to-optimistic Ratio |
|---|---|---|---|
| Env 0 | 1.2501 | 0.9173 | 0.8479 |
| Env 1 | 2.1500 | 0.5809 | 0.4260 |
| Env 2 | 1.0669 | 0.8741 | 0.5527 |
| Env 3 | 1.0784 | 0.7623 | 0.3351 |

*simistic ratio*, *optimistic ratio*, and *closeness-to-optimal ratio*, which offer insights into the framework's effectiveness in navigating partially known environments. The pessimistic ratio compares the actual path length to the cost of an entirely cautious (pessimistic) approach, indicating the efficiency gained by integrating sensing decisions. The optimistic ratio, comparing the actual path length to an idealistic (optimistic) path assuming all unknowns are free, serves as a theoretical benchmark, though it is less reliable in complex environments. The closeness-to-optimal ratio balances these two metrics, showing how well the framework navigates between caution and exploration. The compiled results, shown in Table 5.2, indicate that the framework consistently optimized path length while preserving sensor resources, demonstrating a clear advantage over strategies that rely on fixed or extreme assumptions about unknown areas.

Overall, these findings suggest that the joint sensing and control framework enables robust and adaptable navigation in partially unknown environments. By dynamically adjusting the path based on real-time data and effectively prioritizing high-value sensing opportunities, the framework provides a balanced approach that leverages both exploration and caution to achieve efficient navigation.

## 5.10 Conclusion

The Regret-aware Joint Sensing and Path Planning approach introduced a joint sensing and control framework designed to navigate partially known environments by integrating path planning with sensor activation. Empirical results demonstrated its effectiveness in balancing exploration and efficiency, particularly in simpler environments. The methodology succeeded in dynamically leveraging information gain, regret metrics, and sensor properties to guide the robot's immediate decisions. By focusing on local costs and benefits, the system provided an efficient, reactive solution that reduced unnecessary exploration and enabled the robot to traverse unknown regions with minimal detours.

However, Regret-aware Joint Sensing and Path Planning lacked the ability to plan for long-term sensor allocation. This limitation was most evident in cases where large uncertain areas appeared early in the robot's path. Without a mechanism to anticipate future sensing demands, the system occasionally exhausted its sensor budget prematurely. While this occurred infrequently, it had a significant impact when it did, leaving the robot unable to handle subsequent unknown areas effectively. These failures highlight the method's reliance on immediate cost reductions at the expense of broader path planning considerations.

To address this limitation, future systems must incorporate strategies that anticipate long-term sensor use. By projecting sensing demands across the entire environment and prioritizing high-impact areas, the robot can better conserve resources for critical sections later in its journey. This forward-planning capability forms the foundation for the next methodology, which explicitly integrates sensor use into the planning process. Together, these methodologies illustrate the evolution of sensing and control strategies from reactive to strategic, paving the way for more robust solutions to autonomous navigation in complex environments.

# CHAPTER 6: Methodology 2: Sensor-aware Planner and Regret-based Cost Function

## 6.1 Objective

The primary objective of this methodology is to integrate sensor usage directly into the path-planning process, enhancing the robot's ability to make strategic long-term decisions in partially known environments. Previous methods, such as the one outlined in Chapter 5, provided effective real-time decision-making but lacked the foresight to allocate sensor usage in a way that maximized the efficiency of the overall path. Here, we address this limitation by introducing a global planner that integrates sensor use into the state space, transforming the approach to one that considers sensor activations as an integral component of the path planning itself. This edge-based framework leverages a new state space and cost function, allowing for more deliberate sensor usage that aligns with long-term navigation goals.

## 6.2 Expanded State Space for Sensor-Aware Planning

The expanded state space in this methodology incorporates additional dimensions beyond the conventional environment space, allowing the planner to factor in the implications of sensor usage and the locations of previous sensor activations. This approach is implemented by defining a new State class, where each state encompasses three key attributes:

```
class State:
    pose: Point
    sensor_use: int
    last_activation: Point
```

In this configuration, the pose attribute represents the robot's current position as a Point in the environment. This is analogous to traditional path planning, where the robot's location is a primary component of the state.

The second attribute, *sensor_use*, is an integer that tracks the number of sensor activations used. This enables the planner to account for the limited sensor use budget directly within the state, ensuring that the path planning process respects the constraints on sensor activations. By including *sensor_use*, the system can make informed decisions about future sensor usage based on the remaining budget, thus integrating sensor management as part of the planning process.

Finally, *last_activation* records the location of the most recent sensor activation. This Point serves as a reference for the planner, allowing it to understand the spatial relationship between current movements and past sensing decisions. Tracking *last_activation* ensures that the planner can avoid redundant sensor activations in areas that have already been explored, optimizing sensor allocation by leveraging prior sensing information.

Together, these three attributes—*pose*, *sensor_use*, and *last_activation*—expand the state space into a multidimensional framework that captures the robot's spatial progression and strategic sensor usage. By embedding sensor management within the state definition, the planner can better navigate complex, partially known environments, where strategic sensor allocation is critical for efficient path planning.

## 6.3 Algorithm - Sensor-Aware Multi-dimensional A* Search

The algorithm for this Methodology employs an A* search adapted for Sensor-Aware Planning searching a multi-dimensional space (outlined in Section 6.2), enabling the planner to incorporate sensor constraints into the path planning process. This adaptation of A* addresses the critical need for dynamic sensor management in partially known environments, where efficient sensor use directly impacts navigation efficiency. The core difference between traditional A* and this Sensor-Aware A* algorithm is the integration of checks for sensor usage at each step of neighbor evaluation, ensuring that the path planner adheres to the available sensor budget while optimizing exploration.

### 6.3.1 Core Algorithm Structure

The structure of this Sensor-Aware A* is designed to manage sensor usage dynamically while maintaining the classic A* search methodology for shortest-path planning. In this structure, each state in the search includes attributes for the position, sensor use, and last sensor activation, as previously defined in the State class. Starting from the initial state, each state is evaluated based on its cumulative path cost ($g\_score$) and its estimated total cost to the goal ($f\_score$), prioritizing states with the lowest $f\_score$.

### 6.3.2 Neighbor Evaluation

In evaluating each neighboring cell, the algorithm determines whether exploring that cell would require a new sensor activation. Specifically, it assesses if moving into the neighbor cell would involve transitioning from a known to an unknown region. If the distance between the neighbor and the robot's last sensor activation location ($last\_activation$) is greater than the sensor's radius, a sensor activation is required to proceed.

If moving into the neighboring cell surpasses the remaining sensor budget, that neighbor is excluded from further consideration. This constraint ensures that the path planner evaluates only those paths that remain within the sensor budget, preventing premature exhaustion of sensor resources and preserving them for high-priority exploration.

### 6.3.3 Advantages of Sensor-Aware Planning

One of the primary advantages of this Sensor-Aware A* is its ability to estimate sensor activation points along the path. As each neighbor is evaluated, the algorithm provides an implicit estimate of where the sensor will need to activate in order to proceed, based on the current remaining sensor use $N - N_t$. This feature allows the planner to allocate the sensor resources strategically, avoiding unnecessary sensor activations and reserving the budget for high-value exploration points. By integrating sensor considerations into the path-planning process, this methodology enables more resource-efficient navigation and maximizes the robot's exploratory capabilities in environments with limited sensing capacity.

### 6.3.4 Algorithm Block

The detailed steps of this Sensor-Aware A* are provided in Algorithm 1, where the implementation highlights the neighbor evaluation process and sensor budget checks. This algorithm begins with initializing the start state. It continues by iterating through the open list until the goal is reached or no feasible paths remain within the sensor budget. Through each iteration, it assesses neighbors, updates scores, and adds viable states to the open list. Algorithm 1 showcases the explicit steps in dynamically managing sensor activation decisions as part of the path planning process, illustrating how each component contributes to intelligent sensing in partially known environments.

To illustrate the outputs of the Sensor-Aware A* algorithm, Figure 6.1 depicts the resulting path and the estimated sensor activation positions along that path. The purple dashed line represents the *path*, guiding the robot to the goal, while the *estimated_sensor_positions* are shown as red dots, marking the locations where the algorithm predicts sensors will be activated to gather necessary information.

## 6.4 Edge-Based Cost Function

The edge-based cost function is designed to evaluate the incremental costs between two adjacent cells in the environment: the current cell $x_t$ and the neighboring cell $x_{t+1}$. This function considers the robot's sensor-aware path requirements by assess-

**Algorithm 1** A* with Sensor-Aware Planning (astar_sense)

**Input:** Start position $s$, goal position $g$, remaining sensor uses $sensor\_uses\_remaining$, sensor radius $r_s$, environment $env$, cost function $cost\_function$
Initialize start state $start\_state$ using $s$
Initialize open list $open\_list \leftarrow \{(0, start\_state)\}$ (priority queue)
Initialize dictionaries $came\_from$, $g\_score$, **and** $f\_score$ with initial values for $start\_state$
**while** $open\_list$ is not empty **do**
  $curr \leftarrow$ state with lowest $f\_score$ in $open\_list$
  **if** $curr.pose = g$ **then**
    $final\_state \leftarrow curr$
    Reconstruct $path$ **and** $estimated\_sensor\_positions$ by backtracking $came\_from$
  **end if**
  **for all** neighbors $neighbor$ of $curr.pose$ **do**
    Initialize $prev\_sensor\_pose \leftarrow curr.last\_activation$ **and** $new\_sensor\_uses \leftarrow curr.sensor\_use$
    **if** $neighbor$ is unknown **then**
      **if** $prev\_sensor\_pose$ not within $r_s$ of last activation **or** $curr.pose$ is known **then**
        $prev\_sensor\_pose \leftarrow curr.pose$, increment $new\_sensor\_uses$
      **else**
        $prev\_sensor\_pose \leftarrow None$
      **end if**
    **end if**
    **if** $new\_sensor\_uses > sensor\_uses\_remaining$ **then**
      **continue**
    **end if**
    Calculate $tentative\_g\_score$ using $cost\_function(curr.pose, neighbor)$
    Create $neighbor\_state \leftarrow$ State($neighbor$, $new\_sensor\_uses$, $prev\_sensor\_pose$)
    **if** $neighbor\_state$ is not in $g\_score$ **or** $tentative\_g\_score < g\_score[neighbor\_state]$ **then**
      Update $came\_from$, $g\_score$, **and** $f\_score$ for $neighbor\_state$
      Add $neighbor\_state$ to $open\_list$
    **end if**
  **end for**
**end while**
**Output:** $path$ **and** $estimated\_sensor\_positions$

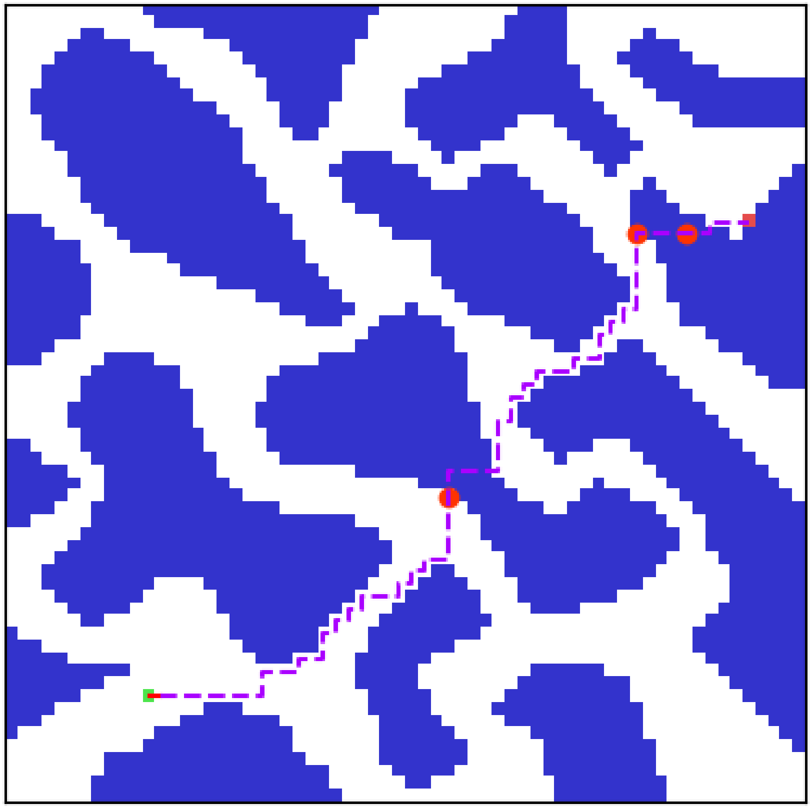


Figure 6.1: Visualization of the Sensor-Aware A* algorithm's outputs, showing the *path* (purple dashed line) and *estimated_sensor_positions* (red dots) along the path.

ing changes in regret and partial pessimistic costs, adapting dynamically based on whether the next cell is known or unknown. By calculating cost differentials, the function directs the robot toward paths that maximize efficiency while strategically managing sensor activations.

The cost calculation for each edge is based on delta formulations that quantify shifts in the robot's information landscape. The first, $\Delta c^{\text{reg}}$, captures the change in regret between the current and next cells. Defined as

$$\Delta c^{\text{reg}} = c^{\text{reg}}(x_{t+1}) - c^{\text{reg}}(x_t), \tag{6.1}$$

this metric allows the planner to assess the potential cost increase associated with exploring or bypassing the unknown, reflecting the cost of moving from a potentially safer route to one with higher uncertainty.

The second key component, the delta in partial pessimistic cost ($\Delta c^{\text{partial\_pess}}$), measures how the cumulative cost of taking an exploratory path changes when transitioning between cells. This delta is computed with an adjustment based on the known or unknown status of $x_t$. If $x_t$ is known, the change is calculated as

$$\Delta c^{\text{partial\_pess}} = \begin{cases} c^{\text{pess}}(x_t) - c^{\text{partial\_pess}}(x_{t+1}), & \text{if } x_t \in \mathcal{F}_t, \\ c^{\text{partial\_pess}}(x_t) - c^{\text{partial\_pess}}(x_{t+1}), & \text{otherwise.} \end{cases} \tag{6.2}$$

This differentiation enables the planner to weigh the benefits of revealing unknown areas against the certainty of following known paths.

The comprehensive edge-based cost function combines these delta values alongside other weighted parameters to produce the final cost for moving between $x_t$ and $x_{t+1}$. As shown in (6.3), the function applies specific weights to each term, incorporating global and local regret values, information gained from sensing, and adjustments for moving into unknown cells. For instance, if the neighboring cell $x_{t+1}$ is an unknown cell ($x_{t+1} \in \mathcal{U}_t$) and sensor activation ($\theta_t = 1$) is triggered, the function includes terms for estimated information gain and the delta in partial pessimistic cost, balanced by the weight parameters $w_4$, $w_5$, and $w_6$. Alternatively, if no sensor activation is required, the transition incurs fewer costs by omitting specific terms associated with sensing.

$$C_t(x_t, x_{t+1}) = \begin{cases} \infty, & \text{if } x_{t+1} \in \mathcal{O}_t \\ w_0, & \text{if } x_{t+1} \in \mathcal{F}_t \ \& \ c^{\text{pess}} = c^{\text{opti}}, \\ w_1 + g^{\text{sense}}(x_{t+1}) \cdot w_2 + \Delta c^{\text{reg}} \cdot w_3, & \text{if } x_{t+1} \in \mathcal{F}_t \ \& \ c^{\text{pess}} \neq c^{\text{opti}}, \\ w_4 + g^{\text{sense}}(x_{t+1}) \cdot w_5 + \Delta c^{\text{partial\_pess}} \cdot w_6 + w_7, & \text{if } x_{t+1} \in \mathcal{U}_t \ \& \ \theta_t = 1, \\ w_4 + g^{\text{sense}}(x_{t+1}) \cdot w_5 + \Delta c^{\text{partial\_pess}} \cdot w_6, & \text{if } x_{t+1} \in \mathcal{U}_t \ \& \ \theta_t = 0. \end{cases} \tag{6.3}$$

Therefore, the edge-based cost function serves as a critical tool in the robot's path planning. Integrating incremental information-based metrics with traditional cost components enables a nuanced approach that optimizes both path efficiency and strategic information acquisition, ensuring the robot's sensor use aligns with long-term navigation goals.

## 6.5 Controller Operation with New Planner

The controller in this methodology operates with a similar framework to the Regret-aware Joint Sensing and Path Planning described in Chapter 5, maintaining an iterative path selection and update process but now utilizing the edge-based cost function and sensor-aware planner. As the robot navigates, it relies on the enhanced cost function to make more refined decisions at each step, considering both the incremental cost of moving between cells and the potential value of sensor activations. When new information is gathered through sensor use, the controller dynamically updates the path, recalculating costs based on the latest data. This adaptive approach ensures that the path remains optimized in real time, reflecting the most current understanding of the environment and allowing the robot to navigate efficiently while conserving sensor resources.

## 6.6 Results

The results from the Sensor-aware Planner and Regret-based Cost Function approach reveal significant improvements in sensor efficiency and path optimization compared to Regret-aware Joint Sensing and Path Planning. The final weights tested in this evaluation, shown in Table 6.1, were selected to effectively balance path length minimization, sensor usage, and prioritization of critical cells based on the edge-based cost function. These weights emphasized areas with high information gain while strategically navigating unknown cells, ensuring that sensor resources were utilized in locations that provided maximal insight into the environment.

Table 6.1: Weight values for the cost function

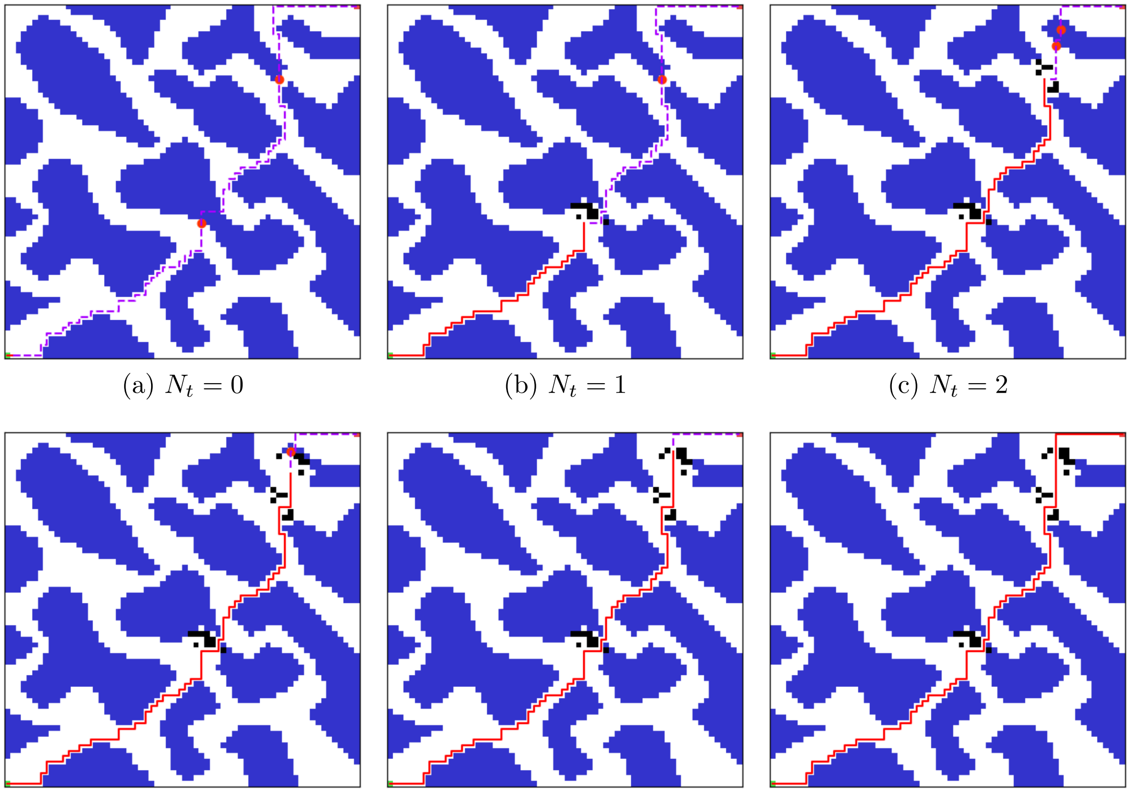
| variable | $w_0$ | $w_1$ | $w_2$ | $w_3$ | $w_4$ | $w_5$ | $w_6$ | $w_7$ |
|---|---|---|---|---|---|---|---|---|
| value | 0.15 | 1.0 | -2.63 | 0.75 | 1.5 | -4.0 | 0.20 | 0.26 |

In the first test environment (see Figure 6.3a), the agent's performance exemplifies the refined path selection process enabled by this new approach. Using the updated cost function, the agent adapts dynamically to the environment, selecting paths that minimize traversal costs while balancing the need for sensor activation. Figure 6.2 shows a sequence of operations within this environment, capturing how the cost map evolves as the robot explores. Each sensor activation recalibrates the map, guiding the robot toward increasingly optimized paths with every update, showcasing the system's real-time adaptability and the advantages of sensor-aware planning.

Across all evaluated environments, the agent's performance metrics—measured through path length ratios and sensor usage—highlight the framework's robustness and adaptability. Averages of these performance ratios across multiple environments are presented in Figure 6.3, demonstrating the scalability of the methodology in varied, complex terrains. Table 6.2 summarizes the quantitative results, with metrics averaged over 50 runs to ensure statistical reliability. This new approach consistently achieved lower path lengths and more strategic sensor activations than prior methods, validating the utility of the integrated edge-based cost function and the expanded search state space in efficiently managing resources while optimizing navigation paths.

Table 6.2: Results averaging the pessimistic ratio, optimistic ratio, and closeness-to-optimistic ratio for Methodology 2.

| | Optimistic Ratio | Pessimistic Ratio | Closeness-to-optimistic Ratio |
|---|---|---|---|
| Env 0 | 122.24% | 92.50% | 70.78% |
| Env 1 | 174.89% | 52.43% | 28.92% |
| Env 2 | 116.82% | 96.82% | 74.65% |
| Env 3 | 113.29% | 75.68% | 33.85% |
| Env 4 | 129.21% | 92.60% | 78.67% |
| Env 5 | 123.91% | 81.87% | 57.13% |



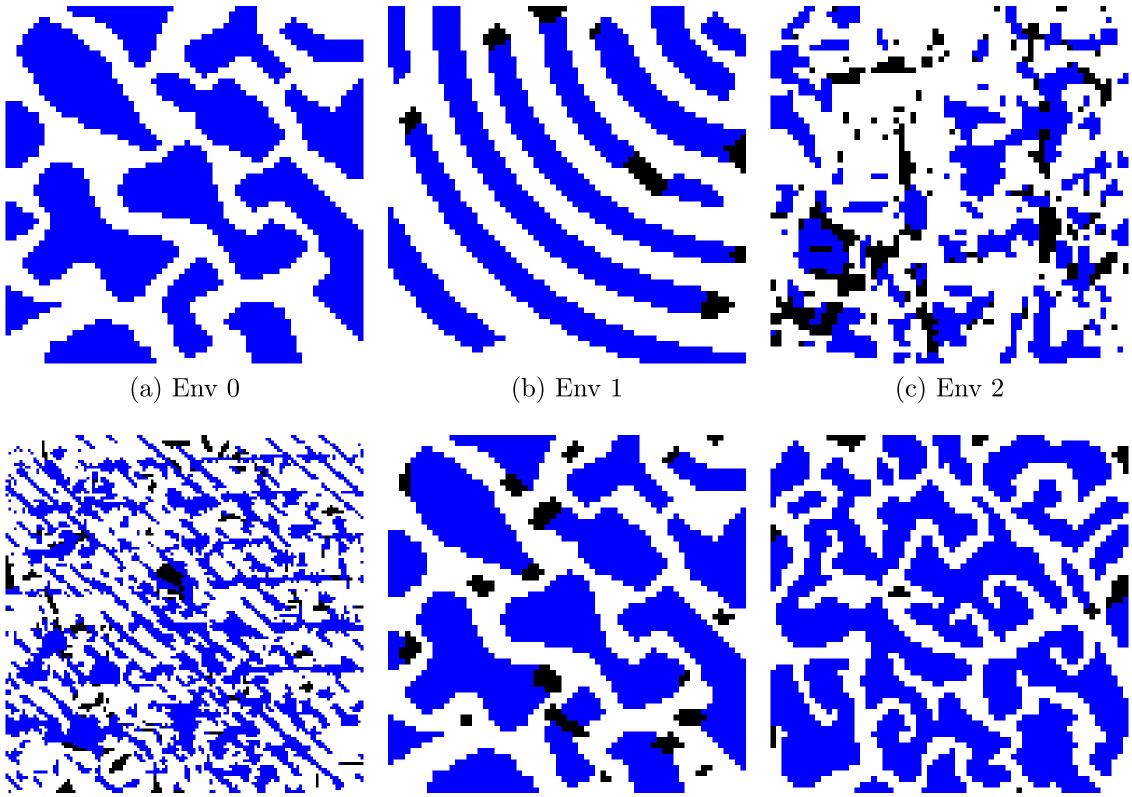
(a) $N_t = 0$ (b) $N_t = 1$ (c) $N_t = 2$

(d) $N_t = 3$ (e) $N_t = 4$ (f) final path

Figure 6.2: Demo of the controller ($N = 4$, $s =$ bottom left in green, $g =$ bottom right in red) navigating the environment with snapshots after every sensor activation. The path is in red, while the proposed path is the purple dashed line. The estimated sensing points are shown in red dots along the path

## 6.7 Conclusion

The *Sensor-aware Planner and Regret-based Cost Function* build on the joint sensing and control approach of the *Regret-aware Joint Sensing and Path Planning* by incorporating long-term sensor planning directly into the path planner's state space. This enhanced approach accounts for not only the robot's position but also its remaining sensor uses and last activation location, expanding the state space from a two-dimensional search to a more complex five-dimensional model. By including sensor usage in the planning algorithm, the *Sensor-aware Planner and Regret-based Cost Function* allows for a more strategic allocation of sensors throughout the robot's journey, helping prevent the early exhaustion of sensing resources.



(a) Env 0 (b) Env 1 (c) Env 2

(d) Env 3 (e) Env 4 (f) Env 5

Figure 6.3: Environments used to evaluate the proposed framework.

One key advantage of this approach is its ability to anticipate and optimize sensor activations based on projected paths, allowing the robot to prioritize high-impact sensing actions and avoid unnecessary activations. The edge-based cost function further enhances this strategic planning by calculating incremental costs for moving between cells, considering changes in regret and partial pessimistic costs. This enables the robot to assess not only the immediate cost of moving into an unknown cell but also the potential value of the information it might gain by doing so, ultimately leading to more efficient navigation through complex environments.

However, the expanded state space also significantly increases computational demands, which can impact real-time performance in larger environments. Future refinements to the Sensor-aware Planner and Regret-based Cost Function could focus on optimizing the computation of the expanded state space or implementing more efficient data structures to manage the increased dimensionality. Despite this limitation, Sensor-aware Planner and Regret-based Cost Function demonstrates a promising advancement in strategic sensor allocation, allowing the robot to conserve resources for crucial decisions along the path and achieve a better balance between exploration and path efficiency.

# CHAPTER 7: Overall Results and Analysis

In this chapter, we present a comparative analysis of the two methodologies—Methodology 1 (*Regret-aware Joint Sensing and Path Planning*) and Methodology 2 (*Sensor-aware Planner and Regret-based Cost Function*). Each approach has its strengths and limitations, particularly regarding sensor usage and path planning strategy, which become evident when deployed in partially unknown environments with limited sensing resources.

## 7.1 Comparison of Sensor Use Strategy

The Regret-aware Joint Sensing and Path Planning methodology excels in balancing sensor activation with path planning in a localized manner, utilizing a joint cost map that prioritizes information gain and short-term path efficiency. This approach, however, lacks a long-term view of sensor usage, which means it may overutilize sensors early in the journey. As a result, in environments where unknown areas are spread over a large space, this can lead to rapid depletion of the sensing budget, making it unsuitable for scenarios requiring long-term sensor management.

The Sensor-aware Planner and Regret-based Cost Function addresses this by integrating the sensor state directly into the planner's state space, effectively embedding a long-term sensor allocation strategy within the path-planning process. By factoring in sensor availability over the entire journey, the Sensor-aware Planner and Regret-based Cost Function methodology allows the robot to conserve sensors for critical regions where unknown areas must be explored. This foresight minimizes early sensor exhaustion and supports more strategic sensing throughout the path. However, this approach has an increased computational cost due to the expanded state space, which can impact real-time performance.

The comparative figures highlight consistent trends across the performance metrics, demonstrating that Methodology 2 typically outperforms or matches Methodology 1 in most environments. For the Optimistic Ratio, Methodology 2 achieves lower or similar ratios, as seen in Figure 7.1, indicating that it is better aligned with the optimistic path strategy, especially in environments with greater complexity. Methodology 1, while occasionally matching Methodology 2 in simpler cases, tends to fall short in environments requiring long-term planning and strategic sensor allocation.

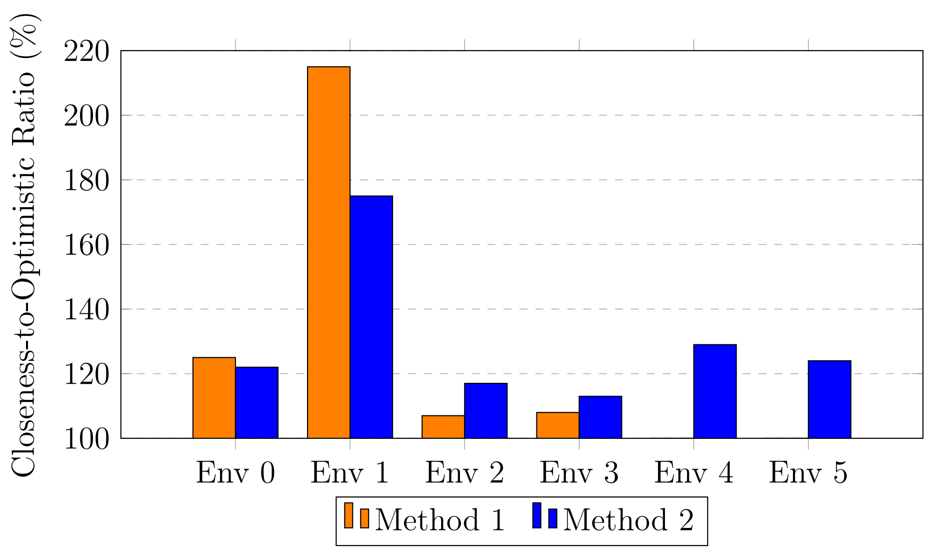



Figure 7.1: Comparison of Optimistic Ratio.

In the Pessimistic Ratio, both methodologies perform similarly across all environments (Figure 7.2), underscoring that both approaches handle pessimistic paths with comparable effectiveness. This suggests that the enhancements in Methodology 2 primarily improve optimistic alignment without compromising cautious navigation.

For the Closeness-to-Optimistic Ratio, Methodology 2 demonstrates a clear advantage in earlier environments, achieving ratios closer to zero and indicating better balance between optimistic and pessimistic strategies (Figure 7.3). However, there are isolated cases where Methodology 1 slightly outperforms Methodology 2, emphasizing the importance of task-specific tuning.

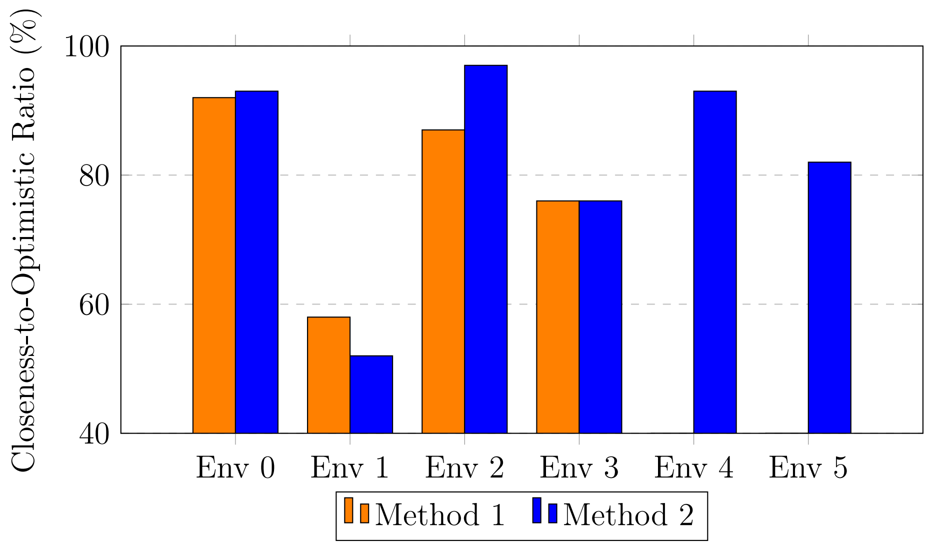



Figure 7.2: Comparison of Pessimistic Ratio.

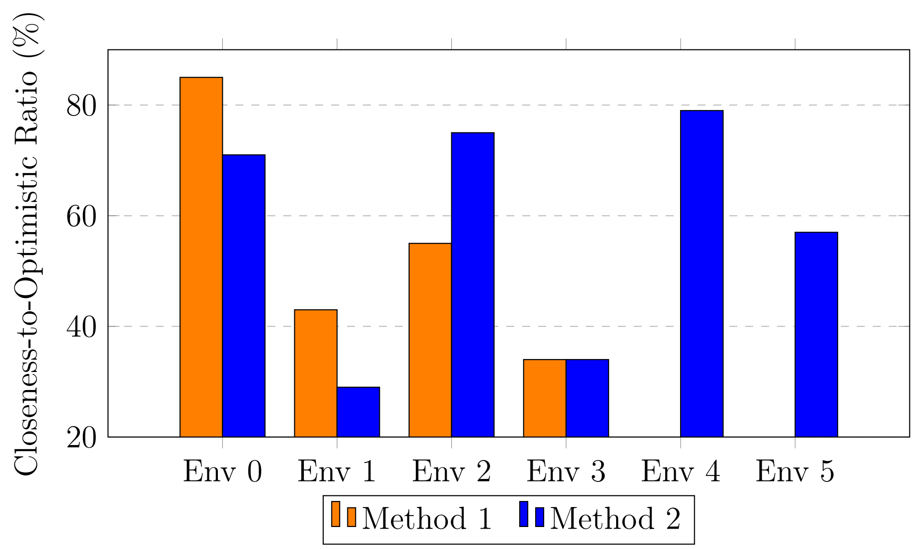
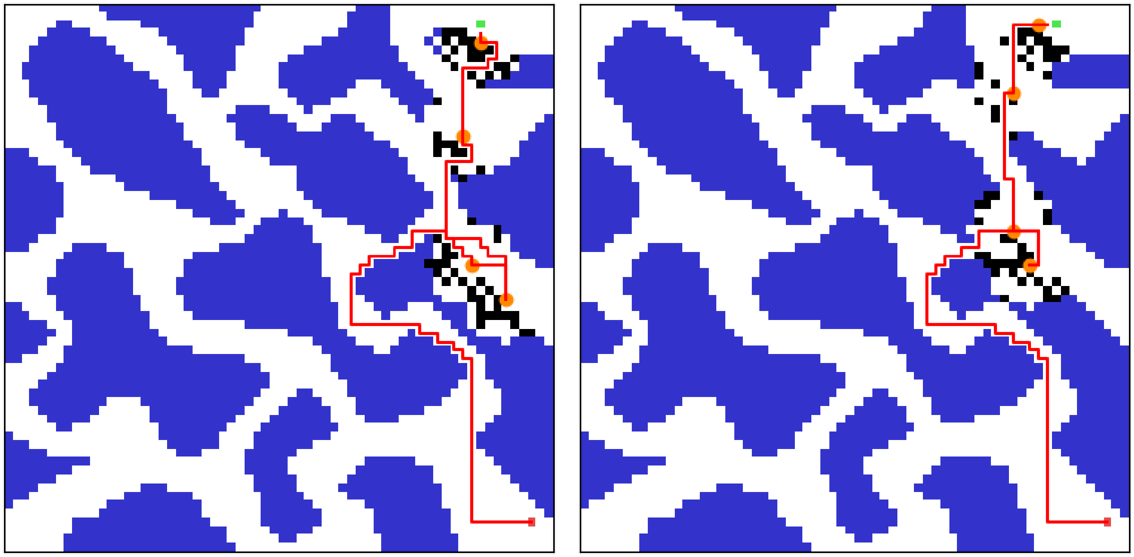



Figure 7.3: Comparison of Closeness-to-Optimistic Ratio.

## 7.2 Comparative Analysis: Method 1 vs. Method 2 in a Shared Environment

To further evaluate the performance of Regret-aware Joint Sensing and Path Planning and Sensor-aware Planner and Regret-based Cost Function, we analyze their respective paths through the same environment with identical starting and goal locations. This direct comparison highlights the practical differences in their sensor use strategies and path-planning efficiency.



(a) Methodology 1 (b) Methodology 2

Figure 7.4: Direct comparison of the 2 methodologies and resulting paths.

As shown in Figure 7.4a, the path generated by Regret-aware Joint Sensing and Path Planning demonstrates its localized optimization strategy. While the robot successfully navigates to the goal, the lack of long-term planning results in suboptimal sensor allocation and a longer overall path. Specifically, the produced a path length of 134 steps. The over-reliance on short-term decision-making and higher sensor activations early in the path are evident from the route taken, which deviates significantly to explore unknown areas unnecessarily.

In contrast, Figure 7.4b illustrates the path generated by Sensor-aware Planner and Regret-based Cost Function, which integrates sensor use into the planning state space. The robot achieves a shorter path length of 116 steps, reflecting more strategic sensor activations and better long-term planning. By conserving sensors for critical points along the path, Sensor-aware Planner and Regret-based Cost Function minimizes unnecessary exploration and ensures efficient navigation toward the goal. This improvement represents a difference of 18 steps and underscores the advantage of incorporating long-term sensor allocation strategies.

The comparison clearly demonstrates that Sensor-aware Planner and Regret-based

Cost Function outperforms Regret-aware Joint Sensing and Path Planning in this scenario, achieving a more efficient path with fewer steps. These results align with broader trends observed in the quantitative metrics, reinforcing the importance of strategic planning in environments with limited sensing resources.

## 7.3 Path Planning Efficiency

The simpler design of Regret-aware Joint Sensing and Path Planning facilitates quicker path calculations, making it suitable for larger environments with limited computational resources, and long-term sensor planning may not be feasible. By focusing on localized information gain and incremental path optimization, it effectively navigates through expansive spaces without extensive computational overhead. This characteristic allows Regret-aware Joint Sensing and Path Planning to handle larger maps with relatively fast processing times, albeit without a strategic long-term sensor allocation approach.

Conversely, Sensor-aware Planner and Regret-based Cost Function's edge-based cost function and expanded state space provide a more comprehensive navigation strategy, optimizing sensor usage across larger environments with multiple unknown regions. This enables the Sensor-aware Planner and Regret-based Cost Function to perform well in scenarios requiring careful sensor allocation and strategic exploration. However, the increased computational requirements of the Method can impact real-time performance, particularly in large-scale environments where resource constraints may become an issue.

## 7.4 Summary of Findings

Both methodologies successfully integrate sensing with path planning in unique ways. Regret-aware Joint Sensing and Path Planning method demonstrates robust performance in environments with fewer unknown regions, where short-term decision-making aligns well with the available sensing resources. On the other hand, the

Sensor-aware Planner and Regret-based Cost Function while computationally heavier, provides a long-term planning capability that supports sensor conservation and efficient navigation across more extensive and complex environments. The choice between the two approaches ultimately depends on the environment's size and the computational resources available, with the method offering a more comprehensive solution where long-term sensor management is critical.

# CHAPTER 8: Conclusions

This thesis presented two methodologies for improving autonomous navigation through joint sensing and control in partially known environments. Both methodologies sought to optimize resource use and path efficiency by integrating path planning with strategic sensor activation. The first methodology (*Regret-aware Joint Sensing and Path Planning*) established a more straightforward and agile approach to navigate unknown regions while balancing sensor activations to minimize path length and avoid unnecessary exploration. The second methodology (*Sensor-aware Planner and Regret-based Cost Function*) built on this foundation by incorporating sensor usage into the state space and applying an edge-based cost function, which allowed for more precise long-term planning of sensor activations. These approaches highlight the potential of combining perception and planning to create more adaptable, efficient navigation strategies in resource-constrained environments.

## 8.1 Summary of Contributions

This thesis aimed to develop efficient joint perception and planning algorithms to navigate partially known environments with explicit sensing constraints. The research presented two primary methodologies: *Regret-aware Joint Sensing and Path Planning*, which emphasized simplicity and computational efficiency, and *Sensor-aware Planner and Regret-based Cost Function*, which introduced a more complex state space for strategic, long-term sensor planning.

*Regret-aware Joint Sensing and Path Planning* created a unified sensing and control framework capable of dynamically balancing path length minimization and intelligent sensor usage in partially unknown terrains. Its simplified design allowed fast path

Table 8.1: Summary and comparison of both methodologies.

| Aspect | Methodology 1: Regret-aware Joint Sensing and Path Planning | Methodology 2: Sensor-aware Planner and Regret-based Cost Function |
|---|---|---|
| **Sensor Use Strategy** | • Localized optimization.<br>• Prioritizes short-term efficiency. | • Long-term sensor allocation.<br>• Strategic sensing for critical areas. |
| **Path Planning Approach** | • Uses joint cost map to control path planning.<br>• Cost function linearly combines weights with various metrics.<br>• Can combine with any path planner. | • Uses custom path planner to plan sensor use.<br>• Edge-based cost function based on change of regret and information gain.<br>• Path planner using expanded state space. |
| **Strengths** | • Fast calculations for real-time use.<br>• Well-suited for environments with sparse unknowns.<br>• Handles large maps efficiently. | • Conserve sensors for critical regions.<br>• Excels in complex and highly uncertain environments.<br>• Provides more strategic and long-term decisions. |
| **Limitations** | • Lacks long-term sensor strategy.<br>• Risk of early sensor depletion with complex uncertainty environments. | • Higher computational cost.<br>• Limited real-time performance in large-scale maps. |
| **Method Contributions** | • Introduces joint cost function that linearly combines regret and VoI based planning<br>• Defines cost based on cell classification, decision uncertainty, and sensing budget.<br>• Enables cost-based trade-offs for path efficiency and information gain. | • Extends the cost function to edge-based calculations, capturing changes in regret and sensor gain between cells.<br>• Integrates sensor activation into the planner state space, creating a dynamic, sensor-aware A*.<br>• Outputs both the path and anticipated sensor positions, enabling resource-efficient navigation. |

calculations suited for more extensive, computationally constrained environments. The joint cost map that prioritized sensor activations enabled the robot to make effective short-term decisions regarding path and sensing. This approach, however, did not consider long-term sensor use, potentially exhausting sensor resources too quickly in dense, unknown regions.

The *Sensor-aware Planner and Regret-based Cost Function* addressed this limitation by integrating sensor use directly into the path planning state space. The method enabled the strategic allocation of sensor resources through an expanded search space and an edge-based cost function, optimizing for longer-term navigation goals. Including edge-based costs, driven by incremental changes in regret and partial pessimistic costs, allowed the robot to evaluate the potential value of sensor activation between cells. This methodology enhanced strategic planning and allowed for more judicious sensor use across the environment. Both methodologies, therefore, contribute to the advancement of autonomous navigation in resource-limited environments, particularly useful for planetary exploration and challenging terrains on Earth.

## 8.2 Limitations

While both methodologies advance joint sensing and path-planning strategies, they also face notable limitations. The expanded state space of Sensor-aware Planner and Regret-based Cost Functionrequires significantly higher computational resources, limiting their applicability in real-time or larger-scale environments. Additionally, both methodologies assume static environments, meaning the robot cannot respond effectively to dynamically changing obstacles or terrain. This reliance on static assumptions constrains the scope of both methods to scenarios with relatively stable, predictable obstacles, such as planetary exploration or pre-mapped disaster zones. However, dynamic environments are typical in real-world scenarios, from shifting terrain to moving obstacles, necessitating more efficient replanning capabilities. Since the implementation is based on A*, it lacks the efficient replanning that algorithms like D* Lite offer, making dynamic updates costly in terms of computation [19].

Another limitation is simplifying the traversal cost, $c(x) = 1$, which was chosen to focus on path length minimization for demonstration purposes. While this simplicity is valuable for analyzing core decision-making strategies, it inherently limits the applicability of the methodologies to real-world scenarios where cost functions may need to account for more complex factors, such as energy consumption, stealth requirements, or terrain difficulty. Adapting $c(x)$ to reflect these nuanced applications is a logical next step, particularly in environments where minimizing detection or conserving energy are mission-critical goals.

Sensor allocation also presents a significant constraint. Although Sensor-aware Planner and Regret-based Cost Function integrates sensor use into the path-planning state, neither methodology accounts for the recharging or replenishing of sensor resources. Robots with a fixed sensor budget are unable to take advantage of intermittent energy sources, such as solar panels or radioisotope thermoelectric generators [63], which could extend mission durations and improve adaptability. The lack of such resource management systems limits the applicability of the methodologies to prolonged missions or highly resource-constrained environments.

Finally, while the regret and cost metrics used in this work provide valuable measures for balancing exploration and caution, they may not capture all aspects of navigation efficiency, particularly in highly stochastic or adversarial environments. For example, environments with unpredictable hazards or adversarial interference might require additional metrics to evaluate safety, robustness, or long-term mission success.

Addressing these limitations would require several enhancements. Incorporating efficient replanning algorithms like D* Lite could allow the robot to respond dynamically to environmental changes. Expanding the definition of $c(x)$ to include multidimensional cost factors could improve the framework's applicability to diverse scenarios. In addition, introducing adaptive resource management strategies, such as forecasting sensor usage or integrating recharging capabilities, would allow the system to operate in more demanding or prolonged missions. By addressing these constraints, the methodologies could evolve into a more comprehensive framework for real-world autonomous navigation.

## 8.3 Future Work

Future research can extend these methodologies in several ways to address the noted limitations and enhance autonomous navigation capabilities. One promising direction is adapting these frameworks to dynamic environments where obstacles or unknown areas change over time. Integrating adaptive path recalculations and real-time cost map updates would allow the robot to respond proactively to environmental shifts, improving reliability in complex, shifting terrains.

Another significant direction involves simulating a rechargeable power system, such as a radioisotope power source or solar power, to manage sensor usage over extended missions. Implementing a rechargeable sensor model could simulate the energy limitations of space missions, requiring the robot to balance energy conservation with recharging cycles. This would allow for intermittent periods of sensor activity, enhancing the robot's capacity to explore while preserving energy.

Extending these methodologies to multi-agent systems offers exciting cooperative exploration and navigation possibilities. Agents could coordinate sensor activations and share information in a multi-agent framework, optimizing path planning across multiple robots. Distributed sensing strategies and collaborative planning would enable teams of robots to map larger areas, reduce redundant sensor usage, and improve resource allocation in expansive or complex environments. These enhancements could pave the way for more sophisticated, cooperative navigation systems that are instrumental in remote exploration or large-scale mapping tasks.

Another promising direction involves extending the methodology to include probabilistic beliefs about the unknown environment, reducing the search space and improving computational efficiency. By incorporating an initial belief about the environment, the system could focus its sensing resources on areas with the highest potential gain. This probabilistic belief framework would refine the sensor-aware A* search, narrowing the regions of interest and significantly reducing computational demands. Additionally, this approach could enhance decision-making by incorporating confidence levels into the path-planning process, ensuring that sensing actions are prioritized where they will significantly impact overall navigation efficiency.

In summary, addressing the current limitations and exploring these avenues of research would significantly broaden the scope and effectiveness of the proposed methodologies, making them more adaptable, scalable, and capable in a wider range of real-world applications.

# REFERENCES


[1] "Google Image Result for off-road robotics." https://images.app.goo.gl/fBPskzKt8QAZjBXA8. [Online; accessed 26-October-2024].

[2] "All Mars Resources - NASA Science — mars.nasa.gov." https://mars.nasa.gov/resources/36238/curiosity-rovers-location-for-sol-630/. [Online; accessed 22-March-2024].

[3] "Google Image Result for search & rescue." https://images.app.goo.gl/kPFq7FGJMQ5TbK526. [Online; accessed 18-November-2024].

[4] A. Wang, "Dynamic Perception Algorithm + IoT – Pudu Elevates Its Technology — linkedin.com." https://www.linkedin.com/pulse/dynamic-perception-algorithm-iot-pudu-elevates-its-amanda-wang-%E7%8E%8B%E8%88%92%E4%BA%88-a3jke/, Oct. 2023. [Online; accessed 17-November-2024].

[5] R. Fearon, "Solar-Powered Robots are the Sustainable Answer to a Cleaner, Connected Planet — discovery.com." https://www.discovery.com/science/solar-powered-robots-are-the-sustainable-answer-to-a-cleaner--co, Aug. 2021. [Online; accessed 18-November-2024].

[6] S. Robotics, "Multi agent robotics systems - SMP Robotics - Autonomous mobile robot." https://smprobotics.com/technology_autonomous_mobile_robot/multi-agent-robotics-systems/. [Online; accessed 18-November-2024].

[7] N. Li, X. Zhao, J. Gao, and X. Cui, "DFA based autonomous decision-making for UGV in unstructured terrain," in *IEEE International Conference on Unmanned Systems*, pp. 34–39, Oct. 2017.

[8] O. Lamarre and J. Kelly, "The importance of adaptive decision-making for autonomous long-range planetary surface mobility," *arXiv preprint arXiv:2409.19455*, Oct. 2024.

[9] Katherine Schauer, "Space Communications: 7 Things You Need to Know - NASA." https://www.nasa.gov/missions/tech-demonstration/space-communications-7-things-you-need-to-know/, Oct. 2020. [Online; accessed 12-September-2023].

[10] T. K. Johnsen and M. Levorato, "Navislim: Adaptive context-aware navigation and sensing via dynamic slimmable networks," in *2024 IEEE/ACM Ninth International Conference on Internet-of-Things Design and Implementation*, pp. 110–121, July 2024.

[11] L. Gerdes, M. Azkarate, J. R. Sánchez-Ibáñez, L. Joudrier, and C. J. Perez-del Pulgar, "Efficient autonomous navigation for planetary rovers with limited resources," *Journal of Field Robotics*, vol. 37, pp. 1153–1170, June 2020.

[12] R. Bajcsy, "Active perception," *Proceedings of the IEEE*, vol. 76, pp. 966–1005, Aug. 1988.

[13] Wikipedia contributors, "Value of information — Wikipedia, the free encyclopedia." `https://en.wikipedia.org/w/index.php?title=Value_of_information&oldid=1186192173`, 2023. [Online; accessed 1-December-2024].

[14] E. W. Dijkstra, "A note on two problems in connexion with graphs," *Numerische Mathematik*, vol. 1, pp. 269–271, Dec. 1959.

[15] P. E. Hart, N. J. Nilsson, and B. Raphael, "A formal basis for the heuristic determination of minimum cost paths," *IEEE Transactions on Systems Science and Cybernetics*, vol. 4, pp. 100–107, July 1968.

[16] A. Stentz, "Optimal and efficient path planning for partially-known environments," in *Proceedings of the 1994 IEEE International Conference on Robotics and Automation*, pp. 3310–3317 vol.4, May 1994.

[17] R. A. Howard, "Information value theory," *IEEE Transactions on Systems Science and Cybernetics*, vol. 2, pp. 22–26, Aug. 1966.

[18] O. Khatib, "Real-time obstacle avoidance for manipulators and mobile robots," in *IEEE International Conference on Robotics and Automation Proceedings*, vol. 2, pp. 500–505, Mar. 1985.

[19] S. Koenig and M. Likhachev, "Fast replanning for navigation in unknown terrain," *IEEE Transactions on Robotics*, vol. 21, pp. 354–363, June 2005.

[20] C. R. V, N. US, N. B. V, S. P, and S. A, "Comparative study of bidirectional A*, D* and D* lite for path planning," in *IEEE International Conference on Electronics, Computing and Communication Technologies*, pp. 1–6, July 2024.

[21] E. Todorov, "General duality between optimal control and estimation," in *47th IEEE Conference on Decision and Control*, pp. 4286–4292, Dec. 2008.

[22] J. D. Pearson, "On the duality between estimation and control," *SIAM Journal on Control*, vol. 4, pp. 594–600, Nov. 1966.

[23] J. S. Baras and A. Bensoussan, "Optimal sensor scheduling in nonlinear filtering of diffusion processes," *SIAM Journal on Control and Optimization*, vol. 27, pp. 786–813, July 1989.

[24] D. Maity, "Optimal intermittent sensing for pursuit-evasion games," *IEEE Control Systems Letters*, vol. 7, pp. 3000–3005, July 2023.

[25] D. Maity and J. S. Baras, "Dynamic, optimal sensor scheduling and value of information," in *18th International Conference on Information Fusion*, pp. 239–244, July 2015.

[26] V. B. Zubek and T. Dietterich, "A POMDP approximation algorithm that anticipates the need to observe," in *Proceedings of the 6th Pacific Rim international conference on Artificial Intelligence*, PRICAI'00, (Berlin, Heidelberg), pp. 521–532, Springer-Verlag, Aug. 2000.

[27] B. Bonet and H. Geffner, "Planning with incomplete information as heuristic search in belief space," in *Proceedings of the Fifth International Conference on Artificial Intelligence Planning Systems*, (Breckenridge, CO, USA), pp. 52–61, AAAI Press, Apr. 2000.

[28] L. Chrisman and R. G. Simmons, "Sensible planning: Focusing perceptual attention," in *American Association for Artificial Intelligence*, pp. 756–761, July 1991.

[29] A. Cassandra, L. Kaelbling, and J. Kurien, "Acting under uncertainty: discrete Bayesian models for mobile-robot navigation," in *Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems*, vol. 2, pp. 963–972 vol.2, Nov. 1996.

[30] H. J. Lim, *Sequential Decision Making under Uncertainty: Optimality Guarantees, Compositional Learning, and Applications to Robotics and Ecology*. PhD thesis, UC Berkeley, 2023.

[31] P. Langley, "Learning to sense selectively in physical domains," in *Proceedings of the first international conference on Autonomous Agents*, pp. 217–226, Feb. 1997.

[32] J. You and W. Wu, "Geometric reinforcement learning based path planning for mobile sensor networks in advection-diffusion field reconstruction," in *IEEE Conference on Decision and Control*, pp. 1949–1954, Dec. 2018.

[33] G. J. Stein, C. Bradley, and N. Roy, "Learning over subgoals for efficient navigation of structured, unknown environments," in *Proceedings of The 2nd Conference on Robot Learning*, pp. 213–222, Oct. 2018.

[34] S. Uppal, A. Agarwal, H. Xiong, K. Shaw, and D. Pathak, "SPIN: Simultaneous perception, interaction and navigation," in *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 18133–18142, June 2024.

[35] B. R. Kiran, I. Sobh, V. Talpaert, P. Mannion, A. A. Al Sallab, S. Yogamani, and P. Pérez, "Deep reinforcement learning for autonomous driving: A survey," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, pp. 4909–4926, June 2021.

[36] J. Delmerico, E. Mueggler, J. Nitsch, and D. Scaramuzza, "Active autonomous aerial exploration for ground robot path planning," *IEEE Robotics and Automation Letters*, vol. 2, pp. 664–671, Apr. 2017.

[37] R. Wang, M. Veloso, and S. Seshan, "Active sensing data collection with autonomous mobile robots," in *IEEE International Conference on Robotics and Automation*, pp. 2583–2588, IEEE, June 2016.

[38] I. Lluvia, E. Lazkano, and A. Ansuategi, "Active mapping and robot exploration: A survey," *IEEE Sensors*, vol. 21, p. 2445, Jan. 2021.

[39] W. Lu, G. Zhang, and S. Ferrari, "An information potential approach to integrated sensor path planning and control," *IEEE Transactions on Robotics*, vol. 30, pp. 919–934, Aug. 2014.

[40] J. Kim, C. D. Eggleton, S. A. Wilkerson, and S. A. Gadsden, "Cooperative sensor-based selective graph exploration strategy for a team of quadrotors," *Journal of Intelligent & Robotic Systems*, vol. 103, p. 24, Sept. 2021.

[41] S. Manjanna and G. Dudek, "Data-driven selective sampling for marine vehicles using multi-scale paths," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 6111–6117, Sept. 2017.

[42] E. A. Hansen, "Cost-effective sensing during plan execution," in *Proceedings of the twelfth national conference on Artificial Intelligence (vol. 2)*, Twelfth National Conference on Artificial Intelligence, (USA), pp. 1029–1035, American Association for Artificial Intelligence, Oct. 1994.

[43] S. Kim and J. Shin, "Optimization-based path planning algorithm with finite traversable area," in *IEEE 26th International Conference on Intelligent Transportation Systems*, pp. 3311–3316, Sept. 2023.

[44] K. H. Low, J. M. Dolan, and P. Khosla, "Active Markov information-theoretic path planning for robotic environmental sensing," *arXiv preprint arXiv:1101.5632*, Jan. 2011.

[45] M. T. J. Spaan, T. S. Veiga, and P. U. Lima, "Decision-theoretic planning under uncertainty with information rewards for active cooperative perception," *Autonomous Agents and Multi-Agent Systems*, vol. 29, pp. 1157–1185, Nov. 2015.

[46] M. Likhachev and A. Stentz, "Information value-driven approach to path clearance with multiple scout robots," in *IEEE International Conference on Robotics and Automation*, pp. 2651–2656, May 2008.

[47] J. Zhao, K. Zhu, S. Li, and X. Yin, "To explore or not to explore: Regret-based ltl planning in partially-known environments," *IFAC-PapersOnLine*, vol. 56, pp. 11337–11343, Jan. 2023.

[48] Y. Zhao, L. Yan, H. Xie, J. Dai, and P. Wei, "Autonomous exploration method for fast unknown environment mapping by using uav equipped with limited fov sensor," *IEEE Transactions on Industrial Electronics*, vol. 71, pp. 4933–4943, May 2024.

[49] Z. Ren, S. Rathinam, M. Likhachev, and H. Choset, "Multi-objective path-based D* lite," *IEEE Robotics and Automation Letters*, vol. 7, pp. 3318–3325, Apr. 2022.

[50] E. Tunstel and A. Howard, "Sensing and perception challenges of planetary surface robotics," in *IEEE SENSORS*, vol. 2, pp. 1696–1701 vol.2, June 2002.

[51] V. Tzoumas, L. Carlone, G. J. Pappas, and A. Jadbabaie, "Sensing-constrained lqg control," in *Annual American Control Conference*, pp. 197–202, June 2018.

[52] A. Singh, A. Krause, C. Guestrin, and W. J. Kaiser, "Efficient informative sensing using multiple robots," *Journal of Artificial Intelligence Research*, vol. 34, pp. 707–755, Apr. 2009.

[53] S. Aggarwal, T. Başar, and D. Maity, "Linear quadratic zero-sum differential games with intermittent and costly sensing," *IEEE Control Systems Letters*, vol. 8, pp. 1601–1606, June 2024.

[54] V. Krishnamurthy, "Decentralized activation in dense sensor networks via global games," *IEEE Transactions on Signal Processing*, vol. 56, pp. 4936–4950, Oct. 2008.

[55] L. Kavraki, P. Svestka, J.-C. Latombe, and M. Overmars, "Probabilistic roadmaps for path planning in high-dimensional configuration spaces," *IEEE Transactions on Robotics and Automation*, vol. 12, pp. 566–580, Aug. 1996.

[56] S. LaValle, "Rapidly-exploring random trees: A new tool for path planning," *Research Report 9811*, 1998.

[57] K. H. Low, J. Dolan, and P. Khosla, "Information-theoretic approach to efficient adaptive path planning for mobile robotic environmental sensing," *Proceedings of the International Conference on Automated Planning and Scheduling*, vol. 19, pp. 233–240, Oct. 2009.

[58] J. Banfi, L. Woo, and M. Campbell, "Is it worth to reason about uncertainty in occupancy grid maps during path planning?," in *International Conference on Robotics and Automation*, pp. 11102–11108, May 2022.

[59] J. Guzzi, R. O. Chavez-Garcia, L. M. Gambardella, and A. Giusti, "On the impact of uncertainty for path planning," in *International Conference on Robotics and Automation*, pp. 5929–5935, May 2019.

[60] D. Conner, A. Rizzi, and H. Choset, "Composition of local potential functions for global robot control and navigation," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, vol. 4, pp. 3546–3551 vol.3, Oct. 2003.

[61] A. I. Panov and K. Yakovlev, "Behavior and path planning for the coalition of cognitive robots in smart relocation tasks," in *Robot Intelligence Technology and Applications 4* (J.-H. Kim, F. Karray, J. Jo, P. Sincak, and H. Myung, eds.), (Cham), pp. 3–20, Springer International Publishing, July 2016.

[62] Wikipedia contributors, "Regret (decision theory) — Wikipedia, the free encyclopedia." `https://en.wikipedia.org/w/index.php?title=Regret_(decision_theory)&oldid=1238005092`, 2024. [Online; accessed 1-December-2024].

[63] Wikipedia contributors, "Radioisotope thermoelectric generator — Wikipedia, the free encyclopedia." `https://en.wikipedia.org/w/index.php?title=Radioisotope_thermoelectric_generator&oldid=1257083691`, 2024. [Online; accessed 1-December-2024].

# APPENDIX A: A* Algorithm

The A* algorithm is a widely used path-planning algorithm that combines features of Dijkstra's algorithm and greedy best-first search to find an optimal path from a start node to a goal node. This appendix provides a concise overview of the algorithm, covering its pseudocode, greedy behavior, and optimality.

## A.1 Algorithm Overview

A* is a graph traversal and path search algorithm that uses a heuristic function to estimate the cost from a given node to the goal. It maintains two primary components:

- **Cost-to-come ($g(n)$)**: The exact cost of the path from the start node to the current node $n$.
- **Cost-to-go ($h(n)$)**: A heuristic estimate of the cost from the current node $n$ to the goal node.

The total estimated cost for each node is given by:

$$f(n) = g(n) + h(n)$$

Nodes are prioritized in a priority queue based on their $f(n)$ values, ensuring the algorithm explores paths that are most promising according to the combined cost.

See Algorithm 2 and Algorithm 3 for algorithm definitions.

## A.2 Greedy Behavior

A* incorporates greedy planning by using the heuristic function $h(n)$, which estimates the cost to reach the goal. When $h(n)$ is highly optimistic, the algorithm behaves like a greedy best-first search, prioritizing nodes that are closer to the goal based on the heuristic alone. While this can lead to faster solutions, it may sacrifice optimality if the heuristic is not admissible.

**Algorithm 2** A* Algorithm

```
open_set ← priority queue containing start with priority 0
came_from ← empty map
g_score ← map with default value ∞
g_score[start] ← 0
f_score ← map with default value ∞
f_score[start] ← heuristic(start, goal)
while open_set is not empty do
    current ← node in open_set with the lowest f_score
    if current = goal then

        return ReconstructPathcame_from, current
    end if
    Remove current from open_set
    for neighbor in neighbors(current, graph) do
        tentative_g_score ← g_score[current] + cost(current, neighbor)
        if tentative_g_score < g_score[neighbor] then
            came_from[neighbor] ← current
            g_score[neighbor] ← tentative_g_score
            f_score[neighbor] ← g_score[neighbor] + heuristic(neighbor, goal)
            if neighbor not in open_set then
                Add neighbor to open_set with priority f_score[neighbor]
            end if
        end if
    end for
end while

return failure {No path found}
```

**Algorithm 3** Reconstruct Path

```
total_path ← current
while current in came_from do
    current ← came_from[current]
    Prepend current to total_path
end while

return total_path
```