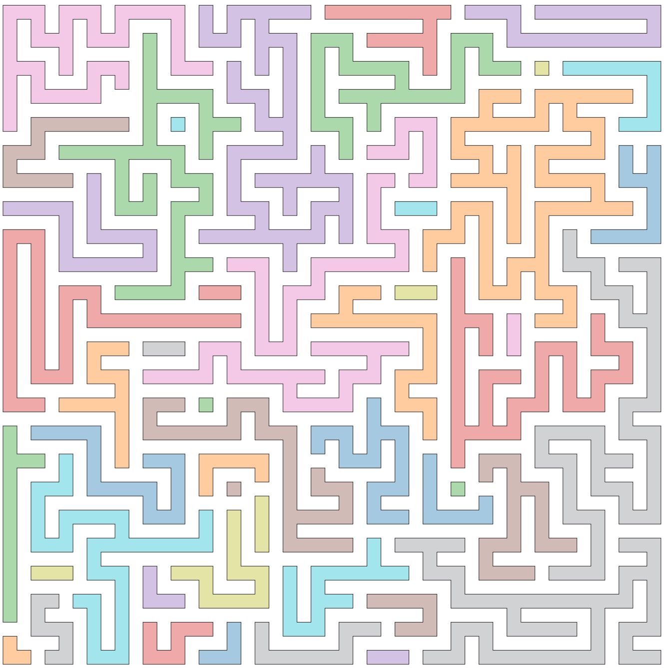
Doctoral Thesis in Computer Science

Exploration and Prediction: Beyond-the-Frontier Autonomous Exploration in Indoor Environments

LUDVIG ERICSON

KTH ROYAL INSTITUTE OF TECHNOLOGY



Exploration and Prediction: Beyond-the-Frontier Autonomous Exploration in Indoor Environments

LUDVIG ERICSON

Academic Dissertation which, with due permission of the KTH Royal Institute of Technology, is submitted for public defence for the Degree of Doctor of Philosophy on Tuesday the 27th of May 2025, at 9:00 a.m. in Kollegiesalen, Brinellvägen 6, Stockholm.

Doctoral Thesis in Computer Science
KTH Royal Institute of Technology
Stockholm, Sweden 2025

TRITA-EECS-AVL-2025:47
ISBN 978-91-8106-270-0

Printed by: Universitetsservice US-AB, Sweden 2025

The struggle itself towards the heights is enough to fill a man's heart. One must imagine Sisyphus happy.

—A. Camus, "The Myth of Sisyphus"

Abstract

Autonomous exploration is a fundamental problem in robotics, where a robot must make decisions about how to navigate and map an unknown environment. While humans rely on prior experience and structural expectations to act under uncertainty, robotic systems typically operate without such priors, exploring reactively based only on what has been observed. The idea of incorporating predictions into exploration has been proposed previously, but the tools required to learn general, high-capacity models have only recently become available through advances in deep learning. This thesis addresses two tightly connected challenges: learning predictive models of indoor environments, and constructing exploration strategies that are able to benefit from such predictions. A core obstacle in this research area is a cyclic dependency: there is little value in developing better predictive models unless exploration methods can make effective use of them, and little value in designing such exploration methods unless reliable models exist. This dependency has historically limited progress. By breaking it, this thesis enables the study and development of both components in tandem. The thesis introduces deep generative models that capture structural regularities in indoor environments using autoregressive sequence modeling. These models outperform traditional approaches in predicting unseen regions beyond the robot's current observations. However, standard exploration methods are shown to perform worse, not better, when informed by accurate predictions. To resolve this, new planning heuristics are proposed, including the distance advantage strategy, which prioritizes exploring regions that are likely to be more difficult to reach in the future. These methods allow predictive models to be used effectively, reducing path length by avoiding situations where the robot must backtrack to previously visited locations. Together, these contributions provide a foundation for autonomous exploration that is informed by learned expectations, and establish a framework where map-predictive modeling and decision-making can be studied and improved jointly.

Sammanfattning

Autonom utforskning är ett grundläggande problem inom robotik, där en robot måste fatta beslut om hur den ska navigera och kartlägga en okänd miljö. Medan människor förlitar sig på tidigare erfarenheter och strukturella förväntningar för att agera under osäkerhet, arbetar robotsystem vanligtvis utan sådana s.k. priors och utforskar reaktivt, enbart baserat på vad som har observerats. Idén att införliva prediktioner i utforskning har föreslagits tidigare, men verktygen som krävs för att lära sig generella, modeller med hög kapacitet har först nyligen blivit tillgängliga genom framsteg inom djupinlärning. Denna avhandling behandlar två nära sammanlänkade utmaningar: att lära sig prediktiva modeller av inomhusmiljöer, och att konstruera utforskningsstrategier som kan dra nytta av sådana prediktioner. Ett centralt hinder inom detta forskningsområde är ett cykliskt beroende: det finns litet värde i att utveckla bättre prediktiva modeller om inte utforskningsmetoder effektivt kan utnyttja dem, och vice versa finns det litet värde i att utforma sådana utforskningsmetoder om inte tillförlitliga modeller finns. Detta beroende har historiskt sett begränsat framsteg. Genom att bryta detta beroende möjliggör denna avhandling parallell utveckling och analys av båda komponenterna. Avhandlingen introducerar djupa generativa modeller som fångar strukturella regelbundenheter i inomhusmiljöer med hjälp av autoregressiv sekvensmodellering. Dessa modeller överträffar traditionella metoder i att förutsäga osedda områden bortom robotens nuvarande observationer. Det visar sig dock att standardmetoder för utforskning presterar sämre, inte bättre, när de informeras av exakta prediktioner. För att lösa detta föreslås nya planeringsheuristiker, inklusive distance advantage-strategin, som prioriterar att utforska områden som sannolikt kommer vara svårare att nå i framtiden. Dessa metoder möjliggör ett effektivt utnyttjande av prediktiva modeller, vilket minskar färdvägens längd genom att undvika situationer där robeten behöver backa tillbaka till tidigare besökta platser, s.k. backtracking. Tillsammans utgör dessa bidrag en grund för autonom utforskning som är informerad av inlärda förväntningar, och etablerar ett ramverk där kartprediktion och beslutsfattande kan studeras och förbättras i samspel.

Preface

The journey to a doctoral degree is a miniature version of the journey through life itself, the hero's journey: to become something greater, we must let go of our old selves. For the doctoral student, every paper is itself a little death. The ideas we believe in and explore exist inside us; they are the thoughts we live and breathe. For one set of ideas to metamorphose into the next, the ego must die and be reborn, making way for something new. It is a painful process, both for grieving the death of what was and for the existential crisis that follows as we reinvent ourselves. But this transformation is not only our own.

No doctoral student works in isolation, we are shaped by those who came before us through our supervisors. Their intellectual legacies, their ambitions, their unrealized projects—these form the invisible scaffolding of our own work. The pressures we feel are not just our own; they are, in part, the weight of what remains undone in the minds of those who guide us. As Jung wrote, *"The greatest burden a child must bear is the unlived life of the parents."* Likewise, the greatest burden a doctoral student must bear is the unlived intellectual life of their supervisor—the lingering research questions of previous students, their unfinished theories, the paths they could not take, now inherited by us. This transfer happens even if, as in my case, the supervisor tries to prevent it. The student, in their crisis of identity, searches for something to cling to, to build their identity from.

A wise man once said that the result of the doctoral program is the person, not the thesis. I always thought this was only superficially true, that the real value is the science done, and that lives in the papers and the results, objective and undeniable. Now that I have written my own thesis I think I see the point. Words on pages are merely evidence of the thoughts and ideas of the people who wrote those words. Sadly, the ideas of the student therefore often wither and die when the student graduates. Be that as it may, I am incredibly grateful for having been given this opportunity to study a subject so deeply. It is an amazing privilege.

This document is a *compilation thesis*, consisting of two parts: the *overview* and the *included papers*. The overview is self-sufficient, a standalone work that comprehensively summarizes the included papers and relates them and their results to the larger field in which they belong. Nevertheless, the papers themselves are part of the thesis, and to understand the thesis, one must understand the papers too.

Acknowledgements

I would like to thank my supervisor Patric for giving me the opportunity to pursue this degree project, for his confidence in me over the years, and for being a mentor to me in all of life's assorted challenges. Your humility, patience, and dedication is truly inspiring. To the wonderful people of my research group, you have been instrumental in keeping the fire burning and coping with the absurdity of it all, I feel blessed to have had your support. 618 truly is not a place, but a state of mind.

To the rock climbers and pebble wrestlers in my life, who knew you could deal with existential crises by climbing to the top of somewhat high rock faces. To the Gotland crew, what an unforgettable time we had, I'll always remember it. Thanks to all the talented, curious, wonderful people at RPL, and to RPL itself. I am honored to have shared in your lives, and I will miss you all incredibly. Thank you to all the students that I had the privilege of teaching throughout the years, nothing is as exciting to me as the vicarious rush of seeing big ideas click into place for the first time.

Tack till alla de som stöttat mig genom åren på KTH, jag kommer aldrig glömma tiden i korridorsboendet på Lappkärrsberget, en livstid utspelade sig där på *Big Brother: Lappis*. Tack till mina språkpartners genom åren, tyvärr blev jag ingen kinesisk poet, men det var värt ett försök.

Ett särskilt tack till David, Isac, Johan, och Jonas, ni har lärt mig så mycket genom de fantastiska personer ni är. Ni besitter den ovanliga men ovärderliga förmågan att värdera och stå upp för er själva och andra, det är sann styrka, och jag inspireras av er varje dag.

Till mina bröder, Ossian och Alexander, jag minns när vi spelade tv-spel tillsammans i vardagsrummet, den slitna röda lädersoffan med insydda knappar, och hur vi fick turas om på familjedatorn som stod på ett alldeles för litet bord, ni har stöpt mig till den jag är idag. Tack mamma, för att du alltid trott på mig och funnits där när det behövts, jag kommer alltid minnas när du höll Felicia i famnen för första gången, då cirkeln slöts och son blev far.

我要感謝斯斯的父母，沒有你們的幫助，我就不能寫這份文件。Jag önskar att jag en dag kan ta lika väl hand om mina barnbarn som ni gör Felicia. Ert stöd har varit ovärderligt.

Framförallt vill jag tacka min familj, Sisi och Felicia. Jag älskar er mer än ord kan beskriva, och ni har gjort mig rikare än pengar någonsin kan. Ingenting gör mig lyckligare än er, och jag är så glad att få dela mitt liv med er.

Ludvig Ericson
March 2025
Stockholm

Contents

Part I

Overview

Chapter 1

Introduction

Humans regularly make decisions despite incomplete knowledge. For example, when Alice is planning to visit Bob tomorrow for dinner, she might not know every detail—what to wear, what route to take, or where to park—but past experience and reasonable assumptions help her navigate these uncertainties successfully. This everyday ability to handle unknowns contrasts sharply with how robots plan their actions, which often requires complete, detailed knowledge of the environment. The first step in practically all robotics tasks is to build a map of the environment which then serves as the single source of truth for the computer program that solves the actual task. This map is constructed from a blank slate, without knowledge of how maps or environments usually look, and without the ability to interpret cues in the environment.

This shortcoming is particularly apparent in the robotics task of *autonomous exploration*, where the goal is for a robot to explore an environment in service of some higher-level goal, e.g., to map it out for downstream tasks, such as when surveying construction sites [1], or to find something, as in search-and-rescue missions in disaster zones [2, 3]. At each moment, the robot decides where next to move by assessing the merit, or *value*, of moving to each of the candidate *frontiers*. Frontiers are places on the boundary of what is known to the robot, and moving towards a frontier pushes the boundary away, increasing the coverage of the environment. The value of a frontier can, in principle, only be determined by computing the optimal tour through the environment, but doing so requires already knowing the environment. This presents a paradox:

> To decide where to explore, the robot must already know what it will find.

In other words, the robot must have *some* expectations about the unknown to make a decision at all, much like Alice. A smarter robot, one that has more accurate expectations, should therefore be better able to evaluate the frontiers, and consequently better execute its task; just as Alice would have made a better choice if Bob had told her that there is no parking near his house.

Indoor environments are highly structured and *quasi-regular*, and a number of basic assumptions can be made—for example: walls are flat and meet at right angles; rooms have at least one human-size door; there can be no rooms outside of the building; and so on. None of these propositions are strictly true, yet they are rarely false either. This quasi-regularity is also found in human language, and it has been famously difficult to construct logic-based computer programs that understand and produce language. Likewise, to date, no programs have shown an understanding of architecture or interior design. Instead, a model must be learned from examples and experience rather than programmed explicitly, and it is only recently that the tools necessary to create such implicit models have become available with the advent of deep learning and large language models. This is the topic of Chapter 3 of this thesis: to model indoor environments using the tools of language models (Paper [A]), and to predict *beyond the frontier* for autonomous exploration (Paper [B]).

More accurate information only improves performance if the robot can use that information effectively, which is not a foregone conclusion. Even if the robot had access to an oracle that told it exactly how the environment will look, computing the optimal exploration tour is computationally unfeasible even in trivial cases due to the extremely large search space, equivalent to the well-known computer science problem the *traveling salesman problem*. Instead, the value of a frontier is approximated as the reward of going there, e.g., how much new information does the robot expect to gain, minus the cost of going there, e.g., how far does it need to travel.

This approximation does not yield better results with better predictions; instead, it has the opposite effect, causing the robot to make worse decisions. To understand why, consider the *corridor-closet gedankenexperiment*, illustrated in Fig. 1.1: the robot is moving along a corridor, and as it comes across a supply closet, it decides between exploring the supply closet, or continuing further down the corridor. The supply closet can only be accessed from this one place and should be explored first; otherwise the robot will eventually have to go back to this place again. If the robot assumes unknown space is simply empty, it will overestimate the reward of the closet and may correctly decide to explore it first. An oracle, on the other hand, would reveal that the closet contains less new information, making the robot erroneously choose the corridor. This is the topic of Chapter 4: how should an algorithm for autonomous exploration be constructed so that it can make use of predictions. Paper [C] first identifies the issue and shows that it is possible to do better with predictions, and Paper [D] presents a practical algorithm that not only is able to use predictions to inform its choices, but performs better than existing exploration methods *without* predictions.

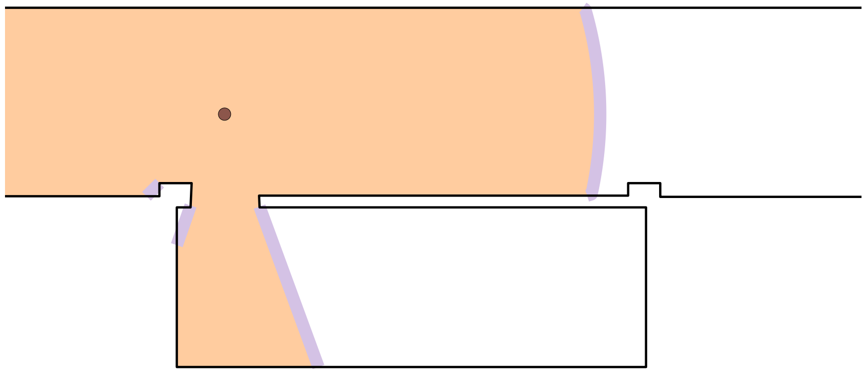


Figure 1.1: The corridor-closet problem is a demonstration of where unnecessary backtracking can occur, i.e., going back over already-explored regions to reach an unexplored region. The robot (brown circle) is facing a choice of either exploring the supply closet below it, or continuing down the corridor to the right. If the robot does not explore the supply closet below as it passes it by, the robot will be forced to go back to this region later. The already-explored region is shaded orange, and the frontier is indicated by a purple outline.

1.1 Scope & Limitations

This thesis aims to study the effects of prior knowledge in autonomous exploration and so seeks to control for other sources of noise and errors. This sets the overall scope of the thesis: the robot must know its location with respect to the environment to build its map, and it must know the map in order to know its location, a classic robotics problem known as *simultaneous localization and mapping* (SLAM). Throughout this thesis, the hypothetical SLAM system is assumed to have no errors and to produce noise-free maps.

In order to reduce the degrees of freedom of the robot, a *holonomic* robot is assumed. A holonomic robot is one that can move in any direction, and this affects what states are easy or difficult to reach: cars are non-holonomic and are difficult to move sideways or rotate on the spot; the gantry of a claw machine, however, is holonomic and can move to any position in any order of movements. A symmetric sensor geometry is also assumed, meaning that the robot has no orientation, only position. The world is considered two-dimensional, except in Paper [C] where it is three-dimensional.

1.2 Contributions & Included Papers

This thesis addresses the following two tightly connected questions:

Q1. How can robots learn to predict unknown regions of the environment to aid autonomous exploration? (Papers [A, B])

Q2. How can better plans be made when prior knowledge is available for autonomous exploration? (Papers [C, D])

Perhaps the most important contribution of this thesis is to break the cyclic dependency between these two questions: there is little reward in solving Q1 until Q2 is solved; likewise, there is little reward in solving Q2 until Q1 is solved. The included papers follow. All papers are available online at the listed DOI or arXiv links.

[A] L. Ericson and P. Jensfelt, "FloorGenT: Generative Vector Graphic Model of Floor Plans for Robotics", in *International Conference on Intelligent Robots and Systems*, IEEE, 2022. DOI: 10.1109/IROS47612.2022.9982144.

[B] L. Ericson and P. Jensfelt, "Beyond the Frontier: Predicting Unseen Walls from Occupancy Grids by Learning from Floor Plans", *IEEE Robotics and Automation Letters*, 2024. DOI: 10.1109/LRA.2024.3410164.

[C] L. Ericson, D. Duberg, and P. Jensfelt, "Understanding Greediness in Map-Predictive Exploration Planning", in *European Conference on Mobile Robots*, IEEE, 2021. DOI: 10.1109/ECMR50962.2021.9568793.

[D] L. Ericson, J. Pedro, and P. Jensfelt, "Information Gain Is Not All You Need", Under review, 2025. arXiv: 2504.01980.

A more comprehensive summary of the included papers is available at the end of this first half of the thesis in Chapter 5. The following is a list of contributed papers that were not included.

[X1] M. C. Welle, L. Ericson, R. Ambruș, and P. Jensfelt, "On the Use of Unmanned Aerial Vehicles for Autonomous Object Modeling", in *European Conference on Mobile Robots*, IEEE, 2017. DOI: 10.1109/ECMR.2017.8098656.

[X2] J. Tang, L. Ericson, J. Folkesson, and P. Jensfelt, "GCNv2: Efficient Correspondence Prediction for Real-Time SLAM", *IEEE Robotics and Automation Letters*, 2019. DOI: 10.1109/LRA.2019.2927954.

[X3] L. Wild, L. Ericson, R. Valencia, and P. Jensfelt, "ExelMap: Explainable Element-based HD-Map Change Detection and Update", in *ECCV Workshop on Vision Centric Autonomous Driving*, Springer, 2024. arXiv: 2409.10178.

Chapter 2

Autonomous Exploration

The beginnings of autonomous exploration lie in *active perception* as proposed by Bajcsy [4] who first identified the connection between perception and exploration when she wrote, *"Perceptual activity is exploratory, probing, searching; percepts do not simply fall onto sensors as rain falls onto ground. We do not just see, we look."* She stresses that perception should not be treated as a passive process but rather one in which the agent deliberately and dynamically controls its sensing parameters. By actively choosing where to look, how to move, and how to focus sensors, the robot reduces its uncertainty of the environment with intent rather than by accident.

The term *autonomous exploration* was first coined by Whaite and Ferrie [5] in their paper "Autonomous Exploration: Driven by Uncertainty". They proposed a framework in which the robot iteratively selects new vantage points to reduce uncertainty in its current model of the environment. By quantifying and actively responding to this uncertainty, the robot determines where and how to move next by maximizing *information gain*. The approach of driving perception by an explicit measure of the agent's knowledge gaps ties closely to the ideas of Bajcsy [4].

Though autonomous exploration did not originally refer to exploration using a mobile platform, that is what the term has come to mean in recent years, and it is how the term will be used in this thesis.

2.1 Problem Definition

Autonomous exploration can be described as the problem of designing an exploration policy $\pi : S \to A$, a mapping from the robot state S to actions A. The robot state is some positional state $x \in X = \mathbb{R}^d$ and the reconstructed map $M \in \mathcal{M} : E \to \{0, 1\}$. The map is an indicator function of occupied points O defined over the domain of *explored* points $E \subseteq X$. The map is constructed from N observations $Z = \{z_1, z_2, \ldots, z_N\} \subseteq \mathcal{Z}$, obtained by the observation function $h : X \to \mathcal{Z}$. The observations Z are intentionally left undefined, noting that they

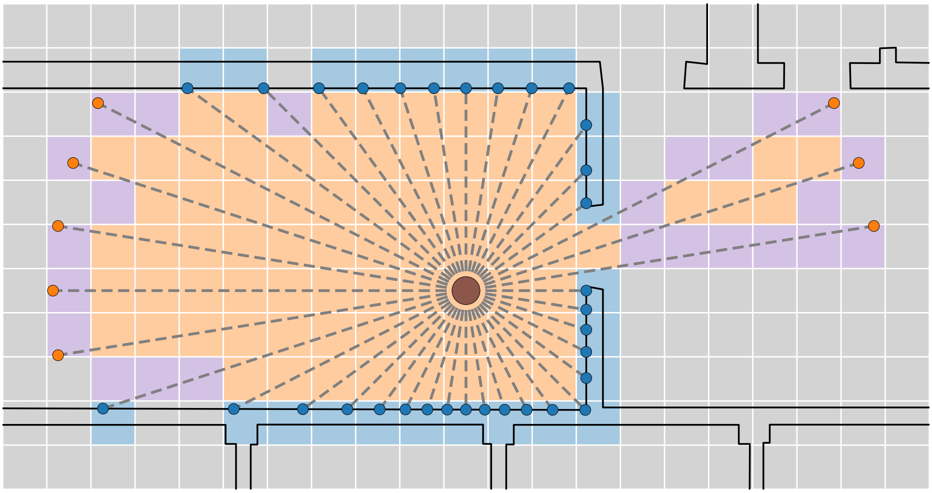


Figure 2.1: Illustration of exploration in the setting of a 2D occupancy grid. The robot (brown circle) is equipped with an omnidirectional range sensor that emits laser pulses, *rays*, in all directions (dashed lines) which measures the time it takes for each ray to reflect back, the *time of flight*, which is proportional to the distance a given ray traveled. Because light decoheres as it propagates, there is an upper range limit; if the ray terminates before then, it is because it is obstructed by the environment, and it is considered a *hit* (blue circles). The occupancy grid cell of each hit is marked occupied (blue); otherwise, where the ray (dashed lines) intersected cells, cells are marked free (orange). Free cells next to unknown cells (gray) are frontier cells (purple). Note that this example sensor only emits a small number of rays whereas real-world range sensors emit hundreds of rays, and the cell sizes have been made larger for the purpose of illustration.

are only relevant to the exploration problem in that they reconstruct the map M. The actions are the possible targets to navigate towards, and is selected among one of the *frontiers* $f \in F$ where the frontier set $F = \partial E \setminus O$ is the unoccupied boundary of explored space. An illustration of the autonomous exploration task is shown in Fig. 2.1. The general approach in exploration is to navigate towards whichever frontier has the maximum score $s : X \to \mathbb{R}$, i.e.,

$$\pi(x, M) = \operatorname*{argmax}_{f \in F} s(f). \tag{2.1}$$

One of the earliest and most influential works in autonomous exploration in the mobile robotics sense is *nearest frontier exploration* by Yamauchi [6], who first introduced the idea of frontiers. As the name suggests, the robot moves towards the nearest frontier, with

$$s(f) = -d_M(x, f) \tag{2.2}$$

and $d_M(x, f)$ is the travel distance from x to f in map M.

The idea of using information gain in autonomous exploration was first proposed by [7] in *Information Based Adaptive Robotic Exploration*, where the key idea is that the exploration policy should consider uncertainty in state estimation since the map is dependent on accurate localization and vice versa, i.e., the SLAM problem. Bourgault et al. [7] thus proposes two terms of information gain; one with respect to state estimation of x, and one with respect to the map M. Subsequent works in autonomous exploration only include the latter term, and the two-term variant has since become *active SLAM.*

The central idea in gain-based exploration policies is to prioritize regions of higher gain first, subject to some penalty for distance. The scoring function is therefore distance d as before, and information gain $I(Z, f)$ at the frontier f,

$$s(f) = \lambda I(Z, h(f)) - d_M(x, f) \tag{2.3}$$

where $\lambda \in \mathbb{R}$ is the *information gain affinity.* A larger λ means the policy will travel longer distances for a given information gain, while $\lambda = 0$ reduces to nearest frontier exploration. Though the definition of I varies between works, the perhaps most ubiquitous class are information-theoretic definitions such as [7], [8], [9], where information gain is the *relative entropy* of the map *belief* distribution $\mathrm{p}(M \mid Z)$ given observations Z and post-observation belief distribution $\mathrm{p}(M \mid Z, z_f)$ after an additional observation $z_f = h(f)$:

$$I(Z, z_f) = \sum_{M \in \mathcal{M}} \mathrm{p}(M \mid Z, z_f) \left(\log \mathrm{p}(M \mid Z, z_f) - \log \mathrm{p}(M \mid Z)\right). \tag{2.4}$$

Intuitively, $I(Z, z_f)$ measures the reduction in entropy of the map belief distribution by an observation z_f. Note that z_f is a random variable for planning purposes, so Eq. (2.4) must be computed in expectation.

So far, the mathematical framework has been kept general to cover all forms of autonomous exploration. The uncountably infinite state space X is however algorithmically challenging, and in practice, it is discretized by either quantization or sampling. In occupancy and voxel grids maps, X is defined as a finite set of *cells* by quantizing with $rx \mapsto \lfloor x \rfloor$, where the resolution r determines the number of cells per unit of length, and $\lfloor \cdot \rfloor$ is the floor operator. The map belief distribution $\mathrm{p}(M \mid Z)$ is then defined as the joint distribution of independent Bernoulli random variables $M_x \sim \mathrm{Bern}(p_x)$ representing the occupancy belief of each cell,

$$\mathrm{p}(M \mid Z) = \prod_{x \in X} \mathrm{p}(M_x = M(x) \mid Z). \tag{2.5}$$

Occupancy belief is initialized to its maximum entropy setting, i.e., $\mathrm{p}(M_x) = \frac{1}{2}$, and is represented in log-odds form with the post-observation belief

$$\log \frac{\mathrm{p}(M_x = 1 \mid Z, z_f)}{\mathrm{p}(M_x = 0 \mid Z, z_f)} = \log \frac{\mathrm{p}(z_f \mid Z, M_x = 1)}{\mathrm{p}(z_f \mid Z, M_x = 0)} + \log \frac{\mathrm{p}(M_x = 1 \mid Z)}{\mathrm{p}(M_x = 0 \mid Z)} \tag{2.6}$$

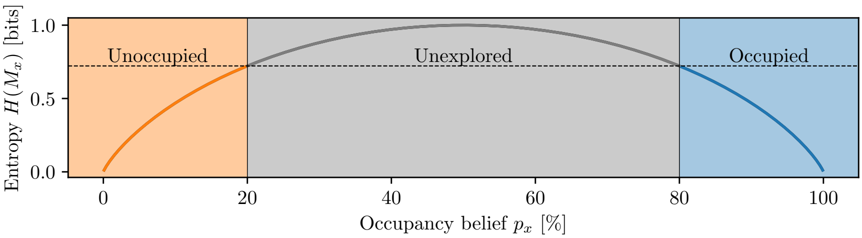


Figure 2.2: The relationship between occupancy belief parameter p_x and the entropy $H(M_x)$ for a single occupancy grid cell. The maximum likelihood estimate $p_x \mapsto [\frac{1}{2} < p_x]$ is only meaningful when the entropy, or uncertainty, is below a predefined threshold (dotted line). Otherwise, the cell is considered unexplored.

Maximizing Eq. (2.4) therefore becomes equivalent to taking actions that maximally push the per-cell occupancy beliefs towards either occupied or unoccupied, illustrated in Fig. 2.2.

It is non-trivial to compute the expectation over all possible observations in Eq. (2.4), and information gain is often approximated by the *belief-of-fact assumption* (BOFA), namely that the maximum likelihood estimate of the map belief is true in explored space, both before and after observation, and that unexplored space has maximum entropy (i.e., one bit). Equation (2.4) then reduces to a single term for the true map M^* after observation,

$$\begin{aligned} I(Z, z_f) &= \log \mathrm{p}(M^* \mid Z, z_f) - \log \mathrm{p}(M^* \mid Z) \\ &= \sum_{x \in X} \left(\log \mathrm{p}(M_x = M_x^* \mid Z, z_f) - \log \mathrm{p}(M_x = M_x^* \mid Z) \right). \end{aligned} \tag{2.7}$$

Since $\mathrm{p}(M_x = M_x^* \mid Z, z_f) = 1$ for cells $x \in X$ explored after observing z_f, but $\mathrm{p}(M_x) = \frac{1}{2}$ for the *novel* cells $x \in F \setminus E$, the sum reduces to

$$I(Z, f) = |F \setminus E| \tag{2.8}$$

measured in bits. In other words, BOFA entails that information gain is equivalent to the number of novel cells observed by visiting a given frontier, which is referred to as *cell-counting gain* in this thesis.

If BOFA holds at least in part, for example because of low-noise sensors, then the map distribution *in unexplored space* determines Eq. (2.4). This is what this thesis will now turn its attention to.

Chapter 3

Modeling Unexplored Space

In principle, the map is not stochastic, and there really is no distribution over maps; the map is the map. The robot, however, can only have *belief* in a map, and its belief is stochastic. In other words, $\mathrm{p}(M \mid Z)$ represents the belief in map M given observations Z. As we saw in the previous chapter, map belief in unexplored space determines information gain under some simple assumptions. This chapter is therefore concerned with modeling map belief in unexplored space, i.e., extrapolating from observations beyond the frontier.

An information-seeking exploration policy will tend to move towards open frontiers regardless of map representation, as unexplored space has maximum entropy by definition. Put differently, it is rarely more informative to remake an observation compared to making a novel observation, except in somewhat contrived circumstances[1]. One of the simplest ways to improve beyond-the-frontier map belief in the occupancy grid representation is to empirically calibrate the unconditional occupancy belief probability $\mathrm{p}(M_x)$; after all, most space is empty. This will cause the policy to be less novelty-seeking, as it now expects to gain less information from novel cells. This is in fact exactly equivalent to attenuating the gain affinity λ in Eq. (2.8), with the relation

$$\lambda \log p_x = \lambda' \log p'_x. \tag{3.1}$$

The first group of works that aim to extrapolate beyond the frontier rely on databases. Ström et al. [10] were first in proposing *predictive autonomous exploration* by matching against a database of previous maps. Oßwald et al. [11] assume a map has been provided, and solve the *offline* autonomous exploration problem where π can be computed ahead of time. In a similar vein, Luperto et al. [12] explore the idea of prior knowledge as given by the user in the form of sketches, bounding boxes, floor plan drawings, or actual maps from previous runs.

[1] If there is a small unexplored hole in the map, it is possible that the information gained from that hole is less than the information gained by observing already-explored space, or more likely, some other open frontier.

Pimentel et al. [13] are perhaps first to predict unexplored space from observations alone, as they propose to use the *Hough line transform* [14] of the occupancy grid to extrapolate the walls into unexplored space on a per-frontier basis, by clustering the frontier into "prediction zones" using a connectivity heuristic, and cell-counting gain is computed for each zone. Luperto et al. [15] predict unexplored space by detecting partially-explored rooms through fitting quadrilaterals to the partially explored map.

Note that in every work reviewed so far, information-seeking exploration policies have been shown to produce substantially *longer* exploration paths than nearest frontier exploration [8], [10], [11], [12], [13]. This counterintuitive result will be put aside for the time being, and instead discussed at length in Chapter 4.

3.1 Learning-based Models

A natural development from explicit approaches are learning-based approaches, which seek to remedy the main drawbacks of explicit approaches: they either require prior knowledge, i.e., there is already a database or map, or the methods have been hand-crafted by experts in an ad-hoc fashion, i.e., line extrapolation and quadrilateral fitting. Learning-based instead model the environment by inductively, learning from experience. Bai et al. [16] use *Gaussian processes* [17], [18] to directly model the information gain $I(Z, f)$. They are perhaps first to point out that there exists an exploration-exploitation tradeoff *within* exploration itself; i.e., should the robot explore so that its model is less uncertain (model exploration), or should it explore what its model believes will improve task performance (model exploitation). In their work, this tradeoff takes the concrete form of exploring frontiers with higher score mean versus those with higher score variance.

Shrestha et al. [19] first proposed to use a *variational auto-encoder* (VAE) as in [20, 21] to learn occupancy beliefs $\mathrm{p}(M \mid Z)$. An auto-encoder is a pair of learned encoder and decoder functions, the encoder $E : \mathcal{M} \to \mathbb{R}^d$ transforms M to a low-dimensional latent code ℓ, the decoder is its inverse $D : \mathbb{R}^d \to \mathcal{M}$, reconstructs M from ℓ. In [19], the two functions are *convolutional neural networks* (CNN) [22, 23] learned by minimizing the reconstruction loss $|\hat{M} - M|$. In a variational auto-encoder, the latent code is a random variable $\ell \sim \mathcal{N}(\mu, \sigma^2)$, and the encoder $E(M) = (\mu, \sigma^2)$ parameterizes the posterior $\mathrm{p}(\ell \mid M)$. The decoder $D(\ell)$ similarly parameterizes the likelihood function $\mathrm{p}(M \mid \ell)$.

Despite multiple attempts by various authors to replicate the results of Shrestha et al. [19], the generated maps have consistently exhibited artifacts such as smudges, holes, and bulges along walls to a degree where the results are unusable. The original network weights were reportedly lost due to a hard drive crash, complicating further replication efforts. However, it may well be that these artifacts stem from inherent limitations of the approach itself. Shrestha et al. [19] deviate from the usual VAE regime by reconstructing the true occupancy grids M^* from the partial occupancy grids M, i.e., not a true auto-encoding formulation. Zangeneh et al.

[24] show that such a formulation is prone to *mode collapse* when learning multi-modal distributions since there may not be any distinctive features in the partial map that allow the encoder to deduce which latent code should be produced. The distribution $p(M^* \mid M)$ is indeed such a multi-modal distribution, since a given partial occupancy M grid can have many valid true maps M^*. Instead, a *conditional* VAE should be used, encoding and decoding the true map M^* in a traditional auto-encoding formulation; the conditioning M is then provided only to the decoder as auxiliary information.

The occupancy grid factorization in Eq. (2.5) assumes independence of individual occupancy beliefs, and drawing samples from $p(M)$ therefore amounts to flipping a biased coin in each cell to decide whether it is occupied, with no regard for neighboring cells. Consider a partially explored room whose depth is unknown, the shape (flat) and orientation (perpendicular) of its back wall are nevertheless nearly certain. Once a single sensor ray establishes the true depth, the uncertainty along that entire wall should ideally vanish. Cell-wise independence precludes such correlations, and the probability mass must be instead be spread among all possible wall positions, leading to the smudging effect in which blurred or averaged walls fail to represent the true structure of the environment, the so-called *mean regression* phenomenon.

Humans excel at this kind of spatial reasoning, yet do not reason in terms of dense, cell-based occupancy grids; instead, we form higher-level abstractions and reason about those. Motivated by this insight, Paper [A] and Paper [B] represent the environment as the line segments of floor plans. Representing walls and other architectural boundaries this way enables capturing the essential geometry, and makes it easier both to generalize across different environments and to reason about structural constraints, ultimately allowing for more robust and accurate models.

3.2 Autoregressive Sequence Modeling

Drawing inspiration from Nash et al. [25], the idea in Paper [A] and Paper [B] is to use the tools of natural language models such as *BERT* [26] and *GPT* [27], [28] to model the geometric relationships in floor plans as a sequence of line segment endpoints. In one-shot methods, it is as if the model is painter that must produce the entire painting in one “stroke”, whereas in the autoregressive regime, that painter now draws individual strokes one at a time, stopping in between each stroke to consider the partially-finished drawing to decide where to start the next stroke.

The key breakthrough in attention-based networks is the *attention* mechanism proposed by Vaswani et al. [29] and its ability to model long-range dependencies in the sequence. Previous deep learning approaches to sequence learning like [30], [31], [32] must propagate information through each step of the sequence, making long-range dependencies difficult. The attention mechanism solves this issue by allowing every token to *attend* to every other token, thus no propagation is necessary. In

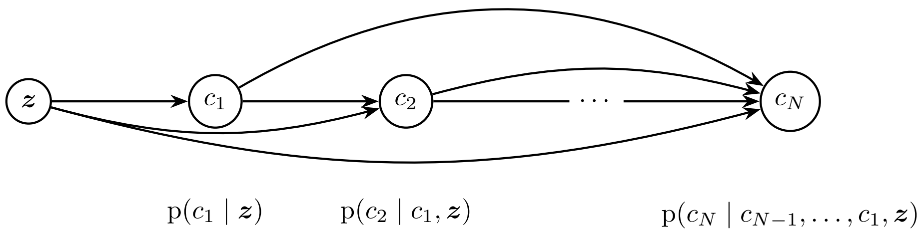


Figure 3.1: Factorization of the joint probability $\mathrm{p}(\boldsymbol{c}, \boldsymbol{z})$ as a probabilistic graphical model. $\boldsymbol{c} = (c_1, c_2, \ldots, c_N)$ is the sequence of tokens, $\boldsymbol{z}$ is conditioning, e.g., sensor data.

generative models, i.e., those that can be sampled, *causal masking* is used so that past tokens cannot attend to future tokens.

The attention mechanism is best understood as a differentiable, associative memory; given a query vector $\boldsymbol{q}$, and the *memory* of key-value pairs $(\boldsymbol{k}_i, \boldsymbol{v}_i)$, the compatibility of each key-query pair is computed by $a_i = \exp(\boldsymbol{q} \cdot \boldsymbol{k}_i)$. The output vector $\boldsymbol{y}$ is the average of the value vectors $\boldsymbol{v}_i$ weighted by a_i. Intuitively, the keys and queries can be thought of as the names of attributes: color, shape, size, etc; the values are the corresponding answers: red, round, large. In *multi-head attention*, each layer consists of many heads, each head able to query different attributes; the next layer then further processes the answers to each head's query.

In autoregressive sequence models, the next-token distribution $p_\theta(c' \mid \boldsymbol{c}, \boldsymbol{z})$ is learned through a deep network $\boldsymbol{y} = f_\theta(\boldsymbol{c}, \boldsymbol{z})$ where $\boldsymbol{y}$ is a parameterization of the probability function, $\boldsymbol{c}$ are the previous tokens, and $\boldsymbol{z}$ is any potential conditioning; a graphical model is provided in Fig. 3.1. In Paper [A] and Paper [B], a categorical distribution is used with $\boldsymbol{y}$ being the *logits* of the next token, transformed to probabilities by the softmax function

$$p_\theta(c' \mid \boldsymbol{c}, \boldsymbol{z}) = \frac{\exp y_{c'}}{\sum_k \exp y_k}. \tag{3.2}$$

The weights θ are found by minimizing the *cross-entropy* between p_θ and the one-hot distribution $q(c') = [c' = \hat{c}']$ of the ground truth next token $\hat{c}'$, which is the same as maximizing the log-likelihood of that token:

$$\begin{aligned} \underset{\theta}{\operatorname{argmin}}\, H(p_\theta, q) &= \underset{\theta}{\operatorname{argmax}} \log p_\theta(\hat{c}' \mid \hat{\boldsymbol{c}}, \boldsymbol{z}) \\ &= \underset{\theta}{\operatorname{argmax}} \left(y_{\hat{c}'} - \text{log-sum-exp}\, \boldsymbol{y} \right). \end{aligned} \tag{3.3}$$

In deep learning, Eq. (3.3) is optimized numerically by some form of *stochastic gradient descent* [33], [34], [35], [36], wherein batches of random samples are drawn uniformly from the set of training pairs

$$\mathcal{D} = \{(c'_1, \boldsymbol{c}_1, \boldsymbol{z}_1), (c'_2, \boldsymbol{c}_2, \boldsymbol{z}_2), \ldots\} \tag{3.4}$$

to approximate the gradient of f_θ. The initial weights are randomly sampled, and a step of length η is taken at each sampling point to obtain the next parameters θ':

$$\theta' = \theta + \eta\,\mathbb{E}\,\nabla_\theta \log p_\theta(\hat{c}' \mid \hat{\boldsymbol{c}}, \boldsymbol{z}). \tag{3.5}$$

The process is then repeated, drawing a new sample of training pairs. Crucially, the attention mechanism is able to process entire sequences at once rather than individual next-token transitions, enabling vastly more efficient training than traditional multi-layer perceptrons without the locality bias of convolutional neural networks.

3.3 Modeling Floor Plans

There is no one-size-fits-all map representation, and modeling maps as they are represented in a given robot embodiment is consequently limited to that embodiment. In Paper [A], we sought to avoid this issue by making a general model over floor plans, allowing users to perform whatever sensor simulation is appropriate for their given embodiment.

A floor plan is a set of line segments L, each line segment is a pair of vertices $\{(a, b), (c, d)\}$. This definition is invariant to the ordering of the segments, as well as the direction of each segment, and an ordering of both line segments and vertices must be established to transform them into canonical sequences. The solution used in both Paper [A] and Paper [B] is a *proximity heuristic*, where line segments are ordered by the distance from the robot to the nearest point on each line segment. The vertices of each segment are then ordered lexicographically. In this way, the set L is unambiguously transformed into an $|L| \times 2 \times 2$ tensor. In Paper [A], this tensor is flattened and the token sequence are the coordinates

$$\boldsymbol{c}(L) = (a_1, b_1, c_1, d_1, \tag{3.6}$$
$$a_2, b_2, c_2, d_2, \ldots) \tag{3.7}$$

as in [25]. However, this means that sample generation is also coordinate-by-coordinate, which is not necessary. Indeed, Paper [A] shows that the uncertainty is lowest for a_i tokens, i.e., the first coordinate of the first vertex, suggesting that the decision is often made already then. For this reason, Paper [B] does not flatten the last dimension of the tensor, the token sequence is thus the vertices themselves

$$\boldsymbol{c}(L) = ((a_1, b_1), (c_1, d_1), \tag{3.8}$$
$$(a_2, b_2), (c_2, d_2), \ldots) \tag{3.9}$$

Finally, the network input $\boldsymbol{x}$ is obtained from $\boldsymbol{c}$ by *discrete embedding*, where each token (i.e., a coordinate or a vertex) is quantized and a look-up table is used to determine the learned embedding vector.

The key difference between Paper [A] and Paper [B] lies in the context $\boldsymbol{z}$. The goal in Paper [A] was to show that language modeling is a viable tool for modeling floor plans for robotics applications, and the context was primarily empty except in the evaluation of rasterization-based generation where the 25 initial line segments were provided as a binary image. The method was evaluated for general geometric indoor understanding, demonstrated by showing that the model could better predict the shortest travel distance to points in the vicinity of the robot from its current location.

The aim in Paper [B], on the other hand, was to extend Paper [A] to a realistic robotics scenario by predicting unseen walls from partial occupancy grids, i.e., precisely the problem faced in autonomous exploration. This is considerably more difficult, as the occupancy grid is the result of past actions, i.e., it is path dependent. For example, a given trajectory can be translated slightly, in whole or in part, to yield a similar but slightly different occupancy grid, and it should produce largely the same output. Paper [B] evaluates the model on predicted versus actual information gain, and it was shown that the line segment-based model outperforms all cell-based approaches evaluated. Crucially, those cell-based approaches all converge to the same error plateau where training error can no longer be reduced, suggesting that the task itself (predicting per-cell occupancy beliefs) is the limiting factor, rather than network architecture or parameter capacity.

3.4 Cluttered Environments

Many map modeling works, e.g., [11], [13], [15], Paper [A], Paper [B], assume that the environment is devoid of *clutter*. Clutter refers to objects in the environment that are typically non-static and small compared to the environment itself, such as furniture, objects on tables, clothes hung on coat racks, etc. Such clutter is not part of the environment *ideal*, but rather temporary guests, and are difficult to model for this reason. This uncluttered assumption is justified for two reasons: first, convenience, as datasets of floor plans such as [37] describe ideals; second, the hypothesis in modeling the map is that the overall environment structure is what matters to exploration, not the clutter in it. For example, it is unrealistic to model what objects will lie on a predicted table in a predicted room. The boundary between clutter and ideal is necessarily ill-defined, but it must be drawn somewhere; floor plans offer such a boundary. It is important to note that exploration can serve many purposes, e.g., in [38], the task is to find an object in an unknown environment; in that case, such objects cannot be considered clutter.

This clutter-free assumption presents a problem for deploying such models in real-world environments which are often cluttered. By an auspicious coincidence, this problem has been at least partially addressed by the technology giants' recent race to provide indoor turn-by-turn navigation. Turn-by-turn navigation faces exactly the same problem of wanting to extract ideal representations from cluttered realities. Apple Inc. [39], for example, provide a toolset for inferring the floor plan directly from a stream of sensor data on their handheld mobile devices. Paper [B] was therefore able to address this limitation by showing that, without adaptation, the model is able to predict unseen parts of the environment from the floor plan recovered by such a toolset.

3.5 Future Work

The problem of map modeling is not limited to autonomous exploration, and a separate strand of similar research has started in the field of autonomous driving. Predicting the road network is in many ways similar to predicting floor plans; they are both planar idealizations of a cluttered 3D world. Liao, Chen, et al. [40] propose to solve this problem in a similar way to Paper [B], by predicting line segments of the road and its lanes. This sparked work on Paper [X3]. Since there is considerably more research effort put into autonomous driving, translating those developments to map prediction is a promising future direction. One could also imagine a combination of the two: autonomous exploration for road networks, the road network models then really are the map models.

More concretely, latent diffusion models [41] are also a promising direction for future work, as these lift one of the main limitations of generative autoregressive sequence models: namely, once a choice is made, i.e., a token is sampled, it is never resampled. Diffusion models instead start from noise, and iteratively refine the entire generated sequence, allowing later positions to influence earlier positions. In fact, a diffusion model can be defined over sets instead of sequences, which is a more natural choice for floor plans as they are ultimately sets of line segments.

A separate but related field that uses floor plans models is found in architecture design tools, or "inverse CAD", i.e., recovering a *computer-aided design* document (CAD) from an image of a floor plan, or sensor data [39], [42], [43], [44], [45]. These works often focus on semantic understanding, which could be interesting as a means of providing robotic intuition in the form of "a fork can be found in the kitchen", which could be part of the pipeline of holistic planning work such as [38]. Recent work using *vision language models* such as [46] and other foundation models have had similar aims, as a means of connecting the semantic to the physical world, such as [47].

As noted previously, improving exploration performance is not as simple as improving the map model, and neither Paper [A] nor Paper [B] were for this reason evaluated as the basis of an exploration policy. This will be the focus of Chapter 4, improving exploration performance with prior map knowledge.

Chapter 4

Predictions and Exploration

As noted in Chapter 1, maximizing information gain does not necessarily improve exploration performance. This has been known since at least 2003, when Stachniss and Burgard [8] showed that *any* gain affinity λ in Eq. (2.3) produces longer paths, shown in Fig. 4.1. They also show the opposite effect for the number of observations, and since their observations are from a sonar, they take a relatively long time to make, and are noisy; λ controls the tradeoff between number of observations and path length. However, as shown in the figure, the path length grows an order of magnitude more than the number of observations shrinks as λ is increased, therefore the potential improvement due to prioritizing information gain is limited.

Indeed, Ström et al. [10] report a factor 1.7 times longer paths for their map-predictive method compared to nearest frontier, arguing that *“The advantages of the prediction-based approach come at a cost—the cost of traversing exploration paths that are longer than the ones generated by the frontier-based approach.”* Oßwald et al. [11] find the optimal exploration tour offline by solving a *traveling salesman problem* (TSP), and find that maximizing information gain produces worse results than nearest frontier. Pimentel et al. [13] show that nearest frontier exploration outperforms both information gain maximization and their proposed method. Interestingly, they also report that time to completion follows the same pattern, disproving the proposition that information gain should be maximized even for the case of slow sensors.

Paper [D] investigates this effect, and we show that *negative* values of λ actually perform best, suggesting, absurdly, that information gain should be *minimized*. The reason is that information gain is a proxy for frontier *depth*: the closet has lower information gain precisely because it is a shallower frontier, i.e., it is easily finished.

That is not to say that there is no place for information gain maximization. In Paper [D], we argue that if the goal is to minimize distance traveled, in what we call *quality-constrained exploration*, then information gain should not be used to score frontiers. Rather, information gain decides which parts of the environment are considered unexplored, i.e., where the frontiers are. If the goal however is to

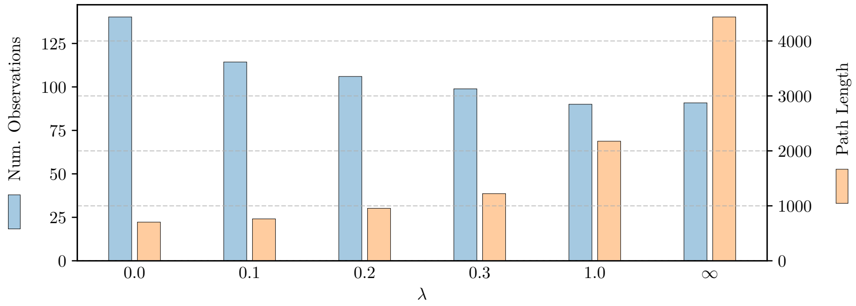


Figure 4.1: Effect of gain affinity λ, adapted from Stachniss and Burgard [8]. Nearest frontier, $\lambda = 0$, produces the least travel distance (path length), but also the largest number of observations. Pure information gain maximization, $\lambda = \infty$, has the opposite effect. The number of observations is reduced by a factor 0.7, while the path length increases by a factor 6.3.

make as few observations as possible, what we call *budget-constrained exploration*—either because of a slow sensor, or a limited resource such as a finite film roll—then it is appropriate to maximize information gain. It is therefore illogical to evaluate a method's travel distance, a quality-constrained behavior, if it is maximizing information gain, a budget-constrained behavior. Instead, a budget-constrained method should be evaluated on the amount of information it collects under its budget. This budget-vs-quality conflation is common in literature, as information gain maximization has become the de-facto standard approach to autonomous exploration. This is the main point of Paper [D]: to challenge the status quo of information gain-driven exploration.

The effect is particularly pronounced when map predictions are used because the information gain can be better maximized and so its consequences become more apparent. It is no surprise then that in evaluating map-predictive methods, authors often choose to terminate exploration before full completion, instead reporting travel distance at some percentage of total *coverage* [15], [19]. Definitions of coverage vary, but generally refers to the volume explored $V(E)$ as a proportion of the volume of all space $V(X)$. Stopping at a percentage of total coverage therefore presupposes $V(X)$ is known *a priori*, and such a *stopping criterion* is consequently impossible to implement in practice. Luperto et al. [48] propose to solve this issue by predicting overall completion via deep learning. We argue instead that it is the conflation of quality-vs-budget constrained exploration that makes it necessary to stop early in the first place. In our framework, quality-constrained exploration stops once there is no more information to gain while minimizing the budget spent; conversely, budget-constrained exploration stops once there is no more budget to spend while maximizing information gained. The choice of paradigm is ultimately application dependent, but one should not evaluate a budget-constrained method with a quality-like metric, and vice versa.

4.1 Beyond Information Gain

It is possible to maximize information gain with respect to the model, and therefore minimizing uncertainty in the distribution of maps. However, the learned distribution of $\mathrm{p}(M^* \mid M)$ is not the same as the empirically-derived $\mathrm{p}(M \mid Z)$, and one must at this point reassert the purpose of autonomous exploration. Even if the model is a perfect oracle, it is not sufficient to produce a predicted map; the environment must be observed directly. A rescue robot that believes it knows the map with low uncertainty cannot terminate its mission early. With that said, there is an exploration-exploitation tradeoff, similar to [16], where the robot can either reduce its model uncertainty so that it can make better plans (explore), or it can execute those better plans to observe the environment (exploit).

Ho et al. [49] recently proposed such a method where the information gain with respect to the map model is maximized, i.e., no exploitation. A stated benefit is that some parts of the map then do not need to be explored since it is certain what they will contain. We argue that this only true if predicting the map is itself the the end goal, i.e., if exploration is done in order to predict a map. We take the opposite standpoint, that the map is predicted in order to do autonomous exploration.

While information-theoretic approaches are no doubt central to the exploration problem, there is something missing in the formulation of Eqs. (2.3) and (2.4). Information gain is computed with respect to a given frontier f, in other words, it only measures the information gain of a single point in space, disregarding the likelihood of obtaining that information elsewhere. This is ultimately why the mistake in the corridor-closet example in Chapter 1 is made, the information gain in the closet is in fact higher if it was computed in expectation over visited future states. This is however intractable to compute, as it is an expectation over future states, observations, and the map belief.

Nearest frontier exploration stands out as one of few exploration paradigms that does not maximize information gain. Still, Yamauchi [6] proposed that nearest frontier exploration is tantamount to doing so: "To gain the most new information about the world, move to the boundary between open space and uncharted territory." No argument is presented to support this proposition, and in fact, the reason that nearest frontier exploration is still used to this day is precisely that it does *not* maximize information gain.

Li et al. [50] showed in 2012 that the margin for improvement compared to the state-of-the-art was small by solving the optimal exploration problem offline, using the tree search method A* [51]. However, the exploration state includes the partially-explored map M, and searching the exploration tree becomes intractable for even trivial map sizes. This is also the main limitation of [50], the map sizes and types are somewhat trivial, and mistakes made by the exploration policy may not be evident in the results because of this. For example, Li et al. [50] report travel distances around 50 m to 100 m, while we in Paper [B] observe distances around 1000 m to 2000 m, roughly a factor of 20 times longer paths. Oßwald et al. [11] further corroborate this analysis by showing that their TSP-based approach in fact significantly reduces travel distance.

One group of works solve TSP online in the partially-explored map [52], [53], seemingly inspired by works performing global path optimization such as [11]. Such a TSP solution answers a different question, namely how to optimally visit all frontiers, however, when the robot approaches a frontier and pushes it back, the TSP solution becomes invalid. The TSP must therefore encode the cost of finishing the frontier, e.g., is it leading into a supply closet, or a corridor. These methods are therefore particularly suited to the map-predictive paradigm, as the TSP can then accurately reflect the cost of finishing each frontier, as opposed to just reaching it.

4.2 Inverse Covisibility Scoring

The main goal of Paper [A] was to determine what kind of planner and predictor are necessary to improve autonomous exploration in the map-predictive paradigm. It is also shown empirically that information gain maximization leads to increased *fragmentation* of unexplored space as small regions of unexplored space are left behind, as quantified by measuring *isoperimetric ratio* (IPR), the ratio of the surface area to the volume.

The proposed approach is called *inverse covisibility scoring* (ICVS), with the goal of prioritizing the exploration of points in space that are the least covisible with other unexplored points. The frontier score is formulated as the integral of score density $\rho(x) \in \mathbb{R}$ over the visible points $h(x)$,

$$s(x, f) = \sum_{x \in h(f) \setminus E} \rho(x) \quad \text{with} \quad \rho(x) = \exp(-\alpha |W(x) \setminus E|). \tag{4.1}$$

The covisible points $W(x) \subseteq X$ is the set of all points y visible from any state s that observes x, i.e., the visible points $h(x)$ of the inverse visibilities $h^{-1}(p)$,

$$W(x) = \{y \in h(s) : s \in h^{-1}(x)\}. \tag{4.2}$$

Note that since $\rho(x)$ tends to one exponentially as the covisible volume $|W(x) \setminus E|$ tends to zero linearly, the score is near-zero in most places except those that observe non-covisible points. The robot is driven into corners and similar, since only unexplored space is considered; the closer the robot gets to the corner, the fewer covisible points there are, and the larger the score grows. A central question in Paper [A] is to assess, once an exploration policy improves with prior map knowledge, quantifying how much prior knowledge is necessary to reach saturation of performance improvement. It is found that already at 1 m to 2 m, performance improves drastically, and saturates at 8 m. The exact numbers are of course contingent on sensor radius, environment size, etc; the point is rather to show that extrapolating from the frontier is a viable approach, rather than having to assume nearly complete prior knowledge as in Oßwald et al. [11] and Li et al. [50].

Paper [A] also evaluates the use of multi-step planning, i.e., not just maximizing the score of a target frontier and planning a shortest path, but instead planning the highest-scoring path directly. The hypothesis was that to solve the corridor-closet issue, the planning depth must be sufficiently large to plan into and then back out of the closet and into the corridor. However, it is shown that such an approach actually confers relatively little benefit over simply choosing the highest-scoring shortest path. This is because multi-step planning is a form of tree search, and even at relatively modest tree depths becomes so large that it is unfeasible to search exhaustively. The fact that the robot should explore the closet before the corridor requires that the robot finds the optimal tour over the entire map, similar to Oßwald et al. [11]. Computing the optimal tour is not feasible in many scenarios for two reasons: first, it requires prior map knowledge to be nearly complete, so that an optimal tour can be planned; and second, it is expensive, and solving the problem offline implies difficulties in replanning if the prior map knowledge turns out to be wrong. ICVS solves both issues, but has at least one critical flaw: it requires sensor simulation when evaluating $h(s)$, and the sensor inverse $h^{-1}(x)$ for essentially every point in E, which is impractical outside of simulated experiments.

4.3 Distance Advantage

The key idea in Paper [D] is that some places are more inaccessible than others, and that the exploration policy should visit generally inaccessible frontiers that are easily accessible from its current location. This is quantified by *distance advantage* which measures the distance to a given frontier $f \in F$ from the current location x, compared to the average distance. More formally, distance advantage scoring is

$$s(x, f) = \mathop{\mathbb{E}}_{y} d_M(y, f) - d_M(x, f). \tag{4.3}$$

The random variable $y \in X$ represents future locations the robot may be at, and plays an important role in determining distance advantage as its distribution function $\mathrm{p}(y)$ is the only design choice. Somewhat problematically, this violates causality as the actual distribution of future states depends on the modeled distribution. One way to escape from this infinite regress is to instead assume that the robot will be moving from some random *source* location, moving by shortest path to some random *target* location. If the distribution of source and target are known, the probability of being at any one state can be computed by *betweenness centrality*, a graph centrality measuring how many shortest paths between pass through each point in the map. $\mathrm{p}(y)$ can then be set to be proportional to how frequently a shortest path passes through y. However, we found that a much simpler approach suffices, letting $\mathrm{p}(y)$ be a uniform distribution over the *local window* around the current location x. The local window is the set of states within some maximum straight-line distance of the robot.

Distance advantage is positive whenever the robot is closer to f than it is expected to be otherwise, and it is zero when the current distance is the average distance; if it is negative, the distance is expected to be lower elsewhere. For example, in the corridor-closet scenario, the robot is closer to the closet than it will be anywhere else, there are no other ways to enter the closet. An example is provided in Fig. 4.2. By contrast, gain-based methods need a conversion factor (i.e., the gain affinity λ) between relative entropy (bits) and travel cost (meters). There is no reason to believe that λ should be a constant, or that the conversion rate should be linear, log-linear, or some other function of gain and cost. Distance advantage compares like-for-like units: length versus length. No proportionality constant is necessary.

Distance advantage is evaluated on its travel distance against nearest frontier and information gain maximization baselines in three environments: a large office space, a cave, and a maze, and we find that nearest frontier produces a factor 1.2 more travel than distance advantage in the office environment, and for information gain maximization, a factor 1.5.

Paper [D] also examines the behavior of distance advantage in the presence of prediction errors, both with false negatives, i.e., failing to predict some parts of the environment, as well as false positives, i.e., predicting non-existent parts of the environment. These errors take the form of clutter, represented by randomly sampling triangles and inserting them into the environment. We found that distance advantage outperforms its baselines even in these cases. Curiously, with extreme cluttering, non-predictive and predictive distance advantage perform almost exactly the same. We reason that this is because the expected distance in Eq. (4.3) is the same in explored and unexplored space, i.e., the path lengths are similarly distributed. In other words, because the distribution is the same with and without predictions, the expected distance is the same, too. Extreme clutter, as opposed to just clutter, is when the environment is substantially different to its original form, blocking off entire regions. Extreme clutter also seems to bring gain-based exploration and nearest frontier exploration to perform the same. This, we reason, happens by a similar mechanism: expected information gain is distributed the same over all frontiers.

Since distance advantage does not involve the sensor model $h(x)$ at all and performs one breadth-first search (BFS) per evaluated frontier f, it solves one of the main drawbacks of Paper [C], namely, its impracticality. Considering a graph of V vertices and E edges, the worst-case time complexity of distance advantage is V instances of BFS, which is $\mathcal{O}(V(V+E))$, and in a four-connected occupancy grid, $E < 4V$. The overall worst-case time complexity is thus $\mathcal{O}(V^2)$.

The main limitation of distance advantage is that it is only applicable in scenarios where path cost is worth optimizing. There are many cases where this may not true, e.g., non-mobile autonomous exploration as proposed by Whaite and Ferrie [5]. The sensor can be moved almost at low cost, so optimizing its path is less important than figuring out and obtaining high-quality views. It is also unclear whether distance advantage is worthwhile for aerial and subnautical scenarios where space is mostly traversible by straight lines.

4.4 Conclusion & Future Work

Paper [D] finally brings to a close the quest for a realistic exploration policy that improves given prior map knowledge. By reasoning from the perspective of prior map knowledge, a novel exploration policy was developed that improves performance even *without* prior map knowledge, representing the first new baseline since 1997 when Yamauchi [6] proposed nearest frontier exploration.

Distance advantage is not a replacement for information gain maximization, but rather a successor to nearest frontier exploration. Instead, information gain maximization should be thought of as an augmentation to nearest frontier exploration, and the same augmentation can be made to distance advantage exploration. Distance advantage does not try to place the sensor optimally; rather, it tries placing the robot itself optimally to reach inaccessible places. This often means staying near walls, which is often at the cost of to sensor coverage.

There are many potential future directions for distance advantage. For example, distance advantage lends itself to multi-agent scenarios, where $\mathrm{p}(y)$ also includes future locations of other robots. Though distance advantage shows improved performance without prior map knowledge, it also improves even more with; it would therefore be interesting to connect distance advantage to a map model such as those proposed in Chapter 3. Equation (4.3) would then be reformulated as an expectation with respect to both y and M.

Another interesting direction is to apply distance advantage subject to a non-trivial motion model, one where the cost of a path is not simply its length. A single-query sampling-based planner such as [54], [55] would likely be too inefficient, and a multi-query representation in the style of *probabilistic roadmaps* [56] would be preferable.

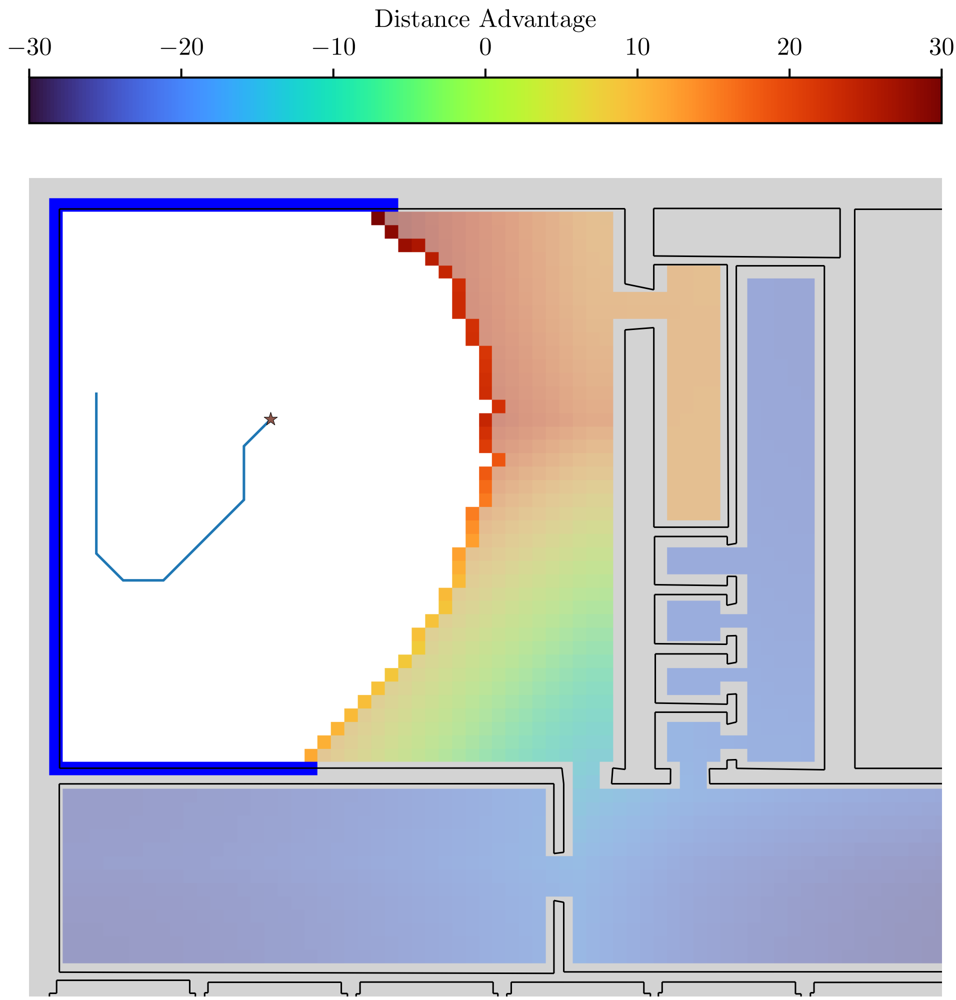


Figure 4.2: Distance advantage at the beginning of exploration with prior map knowledge. Frontier cells F are illustrated as fully-saturated colors; desaturated colors indicate unexplored cells. The scenario is similar to a corridor-closet situation; distance advantage correctly estimates that the closet should be visited first, as it will be further away later on. The reverse is also true, the corridor is further away than it will be in the future. This happens because the corridor connects many rooms, and the distance from most points in each room is lower than this deeper room. The environment is the floor plan of an actual building at a university campus.

Chapter 5

Summary of Included Papers

This chapter contains abstracts of the included papers and contributions by the author of the thesis.

Paper A

FloorGenT: Generative Vector Graphic Model of Floor Plans for Robotics

L. Ericson and P. Jensfelt
In *Proc. Int. Conf. Intelligent Robots and Systems*, 2022

Abstract: Floor plans are the basis of reasoning in and communicating about indoor environments. In this paper, we show that by modelling floor plans as sequences of line segments seen from a particular point of view, recent advances in autoregressive sequence modelling can be leveraged to model and predict floor plans. The line segments are canonicalized and translated to sequence of tokens and an attention-based neural network is used to fit a one-step distribution over next tokens. We fit the network to sequences derived from a set of large-scale floor plans, and demonstrate the capabilities of the model in four scenarios: novel floor plan generation, completion of partially observed floor plans, generation of floor plans from simulated sensor data, and finally, the applicability of a floor plan model in predicting the shortest distance with partial knowledge of the environment.

Contributions by the author: Proposed, designed, and executed ideas and algorithms presented in the paper, and the actual paper.

Paper B

Beyond the Frontier: Predicting Unseen Walls from Occupancy Grids by Learning from Floor Plans

L. Ericson and P. Jensfelt
In *IEEE Robotics and Automation Letters*, 2024

Abstract: In this paper, we tackle the challenge of predicting the unseen walls of a partially observed environment as a set of 2D line segments, conditioned on occupancy grids integrated along the trajectory of a 360° LIDAR sensor. A dataset of such occupancy grids and their corresponding target wall segments is collected by navigating a virtual robot between a set of randomly sampled waypoints in a collection of office-scale floor plans from a university campus. The line segment prediction task is formulated as an autoregressive sequence prediction task, and an attention-based deep network is trained on the dataset. The sequence-based autoregressive formulation is evaluated through predicted information gain, as in frontier-based autonomous exploration, demonstrating significant improvements over both non-predictive estimation and convolution-based image prediction found in the literature. Ablations on key components are evaluated, as well as sensor range and the occupancy grid's metric area. Finally, model generality is validated by predicting walls in a novel floor plan reconstructed on-the-fly in a real-world office environment.

Contributions by the author: Proposed, designed, and executed ideas and algorithms presented in the paper, and wrote the actual paper.

Paper C

Understanding Greediness in Map-Predictive Exploration Planning

L. Ericson, D. Duberg, and P. Jensfelt

In *Euro. Conf. Mobile Robots*, 2021

Abstract: In map-predictive exploration planning, the aim is to exploit a-priori map information to improve planning for exploration in otherwise unknown environments. The use of map predictions in exploration planning leads to exacerbated greediness, as map predictions allow the planner to defer exploring parts of the environment that have low value, e.g., unfinished corners. This behavior is undesirable, as it leaves holes in the explored space by design. To this end, we propose a scoring function based on inverse covisibility that rewards visiting these low-value parts, resulting in a more cohesive exploration process, and preventing excessive greediness in a map-predictive setting. We examine the behavior of a non-greedy map-predictive planner in a bare-bones simulator, and answer two principal questions: a) how far beyond explored space should a map predictor predict to aid exploration, i.e., is more better; and b) does shortest-path search as the basis for planning, a popular choice, cause greediness. Finally, we show that by thresholding covisibility, the user can trade-off greediness for improved early exploration performance.

Contributions by the author: Formulated the problem together with D. Duberg, designed and implemented the solution presented in the paper, and wrote the actual paper.

Paper D

Information Gain Is Not All You Need

L. Ericson, J. Pedro, and P. Jensfelt
Under review

Abstract: Autonomous exploration in mobile robotics is driven by two competing objectives: coverage, to exhaustively observe the environment; and path length, to do so with the shortest path possible. Though it is difficult to evaluate the best course of action without knowing the unknown, the unknown can often be understood through models, maps, or common sense. However, previous work has shown that improving estimates of information gain through such prior knowledge leads to greedy behavior and ultimately causes backtracking, which degrades coverage performance. In fact, any information gain maximization will exhibit this behavior, even without prior knowledge. Information gained at task completion is constant, and cannot be maximized for. It is therefore an unsuitable choice as an optimization objective. Instead, information gain is a decision criterion for determining which candidate states should still be considered for exploration. The task therefore becomes to reach completion with the shortest total path. Since determining the shortest path is typically intractable, it is necessary to rely on a heuristic or estimate to identify candidate states that minimize the total path length. To address this, we propose a heuristic that reduces backtracking by preferring candidate states that are close to the robot, but far away from other candidate states. We evaluate the performance of the proposed heuristic in simulation against an information gain-based approach and frontier exploration, and show that our method significantly decreases total path length, both with and without prior knowledge of the environment.

Contributions by the author: The paper has shared authorship between L. Ericson and J. Pedro. Formulated the problem and co-developed solution together with J. Pedro, implemented and executed method and evaluation, co-authored the paper with J. Pedro, with special focus on producing figures and experimental results.

References

[A] L. Ericson and P. Jensfelt, "FloorGenT: Generative Vector Graphic Model of Floor Plans for Robotics", in *International Conference on Intelligent Robots and Systems*, IEEE, 2022. DOI: 10.1109/IROS47612.2022.9982144.

[B] L. Ericson and P. Jensfelt, "Beyond the Frontier: Predicting Unseen Walls from Occupancy Grids by Learning from Floor Plans", *IEEE Robotics and Automation Letters*, 2024. DOI: 10.1109/LRA.2024.3410164.

[C] L. Ericson, D. Duberg, and P. Jensfelt, "Understanding Greediness in Map-Predictive Exploration Planning", in *European Conference on Mobile Robots*, IEEE, 2021. DOI: 10.1109/ECMR50962.2021.9568793.

[D] L. Ericson, J. Pedro, and P. Jensfelt, "Information Gain Is Not All You Need", Under review, 2025. arXiv: 2504.01980.

[X1] M. C. Welle, L. Ericson, R. Ambruș, and P. Jensfelt, "On the Use of Unmanned Aerial Vehicles for Autonomous Object Modeling", in *European Conference on Mobile Robots*, IEEE, 2017. DOI: 10.1109/ECMR.2017.8098656.

[X2] J. Tang, L. Ericson, J. Folkesson, and P. Jensfelt, "GCNv2: Efficient Correspondence Prediction for Real-Time SLAM", *IEEE Robotics and Automation Letters*, 2019. DOI: 10.1109/LRA.2019.2927954.

[X3] L. Wild, L. Ericson, R. Valencia, and P. Jensfelt, "ExelMap: Explainable Element-based HD-Map Change Detection and Update", in *ECCV Workshop on Vision Centric Autonomous Driving*, Springer, 2024. arXiv: 2409.10178.

[1] S. Omari, P. Gohl, M. Burri, M. Achtelik, and R. Siegwart, "Visual Industrial Inspection using Aerial Robots", in *International Conference on Applied Robotics for the Power Industry*, IEEE, 2014.

[2] D. Calisi, A. Farinelli, L. Iocchi, and D. Nardi, "Autonomous Exploration for Search and Rescue Robots", *Transactions on the Built Environment*, 2007.

[3] F. Colas, S. Mahesh, F. Pomerleau, M. Liu, and R. Siegwart, "3D Path Planning and Execution for Search and Rescue Ground Robots", in *International Conference on Intelligent Robots and Systems*, IEEE, 2013.

[4] R. Bajcsy, "Active Perception", *Proceedings of the IEEE*, vol. 76, no. 8, pp. 966–1005, 1988.

[5] P. Whaite and F. P. Ferrie, “Autonomous Exploration: Driven by Uncertainty”, *Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 3, pp. 193–205, 1997.

[6] B. Yamauchi, “A Frontier-based Approach for Autonomous Exploration”, in *International Symposium on Computational Intelligence in Robotics and Automation*, IEEE, 1997.

[7] F. Bourgault, A. A. Makarenko, S. B. Williams, B. Grocholsky, and H. F. Durrant-Whyte, “Information Based Adaptive Robotic Exploration”, in *International Conference on Intelligent Robots and Systems*, IEEE, 2002.

[8] C. Stachniss and W. Burgard, “Exploring Unknown Environments with Mobile Robots using Coverage Maps”, in *International Joint Conference on Artificial Intelligence*, 2003.

[9] M. Selin, M. Tiger, D. Duberg, F. Heintz, and P. Jensfelt, “Efficient Autonomous Exploration Planning of Large-Scale 3D Environments”, *Robotics and Automation Letters*, 2019.

[10] D. P. Ström, F. Nenci, and C. Stachniss, “Predictive Exploration Considering Previously Mapped Environments”, in *International Conference on Robotics and Automation*, IEEE, 2015.

[11] S. Oßwald, M. Bennewitz, W. Burgard, and C. Stachniss, “Speeding-up Robot Exploration by Exploiting Background Information”, *Robotics and Automation Letters*, vol. 1, no. 2, pp. 716–723, 2016.

[12] M. Luperto, M. Antonazzi, F. Amigoni, and N. A. Borghese, “Robot Exploration of Indoor Environments using Incomplete and Inaccurate Prior Knowledge”, *Robotics and Autonomous Systems*, 2020.

[13] J. M. Pimentel, M. S. Alvim, M. F. Campos, and D. G. Macharet, “Information-Driven Rapidly-Exploring Random Tree for Efficient Environment Exploration”, *Journal of Intelligent & Robotic Systems*, vol. 91, pp. 313–331, 2018.

[14] R. O. Duda and P. E. Hart, “Use of the Hough Transformation to Detect Lines and Curves in Pictures”, *Communications of the ACM*, vol. 15, no. 1, pp. 11–15, 1972.

[15] M. Luperto, V. Arcerito, and F. Amigoni, “Predicting the Layout of Partially Observed Rooms from Grid Maps”, in *International Conference on Robotics and Automation*, IEEE, 2019.

[16] S. Bai, J. Wang, F. Chen, and B. Englot, “Information-theoretic Exploration with Bayesian Optimization”, in *International Conference on Intelligent Robots and Systems*, IEEE, 2016.

[17] C. E. Rasmussen, “Gaussian Processes in Machine Learning”, in *Summer School on Machine Learning*, Springer, 2003, pp. 63–71.

[18] S. O'Callaghan, F. Ramos, and H. Durrant-Whyte, "Contextual Occupancy Maps using Gaussian processes", in *International Conference on Robotics and Automation*, IEEE, 2009, pp. 1054–1060.

[19] R. Shrestha, F.-P. Tian, W. Feng, P. Tan, and R. Vaughan, "Learned Map Prediction for Enhanced Mobile Robot Exploration", in *International Conference on Robotics and Automation*, IEEE, 2019.

[20] D. P. Kingma, M. Welling, et al., "Auto-Encoding Variational Bayes", in *International Conference on Learning Representations*, Banff, Canada, 2014.

[21] D. J. Rezende, S. Mohamed, and D. Wierstra, "Stochastic Backpropagation and Approximate Inference in Deep Generative Models", in *International Conference on Machine Learning*, PMLR, 2014.

[22] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based Learning Applied to Document Recognition", *Proceedings of the IEEE*, 1998.

[23] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks", *Advances in Neural Information Processing Systems*, 2012.

[24] F. Zangeneh, L. Bruns, A. Dekel, A. Pieropan, and P. Jensfelt, "A Probabilistic Framework for Visual Localization in Ambiguous Scenes", in *International Conference on Robotics and Automation*, IEEE, 2023.

[25] C. Nash, Y. Ganin, S. A. Eslami, and P. Battaglia, "Polygen: An autoregressive generative model of 3D meshes", in *International Conference on Machine Learning*, PMLR, 2020.

[26] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "Bert: Pre-training of Deep Bidirectional Transformers for Language Understanding", 2018. arXiv: `1810.04805`.

[27] A. Radford, J. Wu, R. Child, D. Luan, D. Amodei, I. Sutskever, et al., "Language models are unsupervised multitask learners", *OpenAI blog*, 2019. [Online]. Available: `https://openai.com/blog/better-language-models/`.

[28] T. Brown, B. Mann, N. Ryder, M. Subbiah, J. D. Kaplan, P. Dhariwal, A. Neelakantan, P. Shyam, G. Sastry, A. Askell, S. Agarwal, A. Herbert-Voss, G. Krueger, T. Henighan, R. Child, A. Ramesh, D. Ziegler, J. Wu, C. Winter, C. Hesse, M. Chen, E. Sigler, M. Litwin, S. Gray, B. Chess, J. Clark, C. Berner, S. McCandlish, A. Radford, I. Sutskever, and D. Amodei, "Language Models are Few-Shot Learners", in *Advances in Neural Information Processing Systems*, Curran Associates, Inc., 2020.

[29] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin, "Attention Is All You Need", in *Advances in Neural Information Processing Systems*, Curran Associates, Inc., 2017.

[30] J. L. Elman, "Finding Structure in Time", *Cognitive Science*, vol. 14, no. 2, pp. 179–211, 1990.

[31] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory", *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[32] A. Graves, A.-r. Mohamed, and G. Hinton, "Speech Recognition with Deep Recurrent Neural Networks", in *International Conference on Acoustics, Speech and Signal Processing*, IEEE, 2013, pp. 6645–6649.

[33] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning Representations by Back-propagating Errors", *nature*, 1986.

[34] L. Bottou, "Online Learning and Stochastic Approximations", *Online Learning in Neural Networks*, vol. 17, no. 9, p. 142, 1998.

[35] D. P. Kingma and J. Ba, "Adam: A Method for Stochastic Optimization", in *International Conference on Learning Representations*, 2015.

[36] T. Bachlechner, B. P. Majumder, H. Mao, G. Cottrell, and J. McAuley, "ReZero is All You Need: Fast Convergence at Large Depth", in *Uncertainty in Artificial Intelligence*, PMLR, 2021.

[37] A. Aydemir, P. Jensfelt, and J. Folkesson, "What can we learn from 38 000 rooms? Reasoning about unexplored space in indoor environments", in *Conference on Intelligent Robots and Systems*, IEEE, 2012.

[38] N. Hawes, C. Burbridge, F. Jovan, L. Kunze, B. Lacerda, L. Mudrová, J. Young, J. Wyatt, D. Hebesberger, T. Körtner, R. Ambrus, N. Bore, J. Folkesson, P. Jensfelt, L. Beyer, A. Hermans, B. Leibe, A. Aldoma, T. Fäulhammer, M. Zillich, M. Vincze, E. Chinellato, M. Al-Omari, P. Duckworth, Y. Gatsoulis, D. C. Hogg, A. G. Cohn, C. Dondrup, J. P. Fentanes, T. Krajník, J. M. Santos, T. Duckett, and M. Hanheide, "The STRANDS project: Long-term Autonomy in Everyday Environments", *Robotics & Automation Magazine*, vol. 24, no. 3, pp. 146–156, 2017. DOI: `10.1109/MRA.2016.2636359`.

[39] Apple Inc., *3D Parametric Room Representation with RoomPlan*, 2022. [Online]. Available: `https://machinelearning.apple.com/research/roomplan`.

[40] B. Liao, S. Chen, et al., "MapTR: Structured Modeling and Learning for Online Vectorized HD Map Construction", in *International Conference on Learning Representations*, 2023.

[41] R. Rombach, A. Blattmann, D. Lorenz, P. Esser, and B. Ommer, "High-Resolution Image Synthesis with Latent Diffusion Models", in *Computer Vision and Pattern Recognition*, 2022.

[42] Y. Yue, T. Kontogianni, K. Schindler, and F. Engelmann, "Connecting the Dots: Floorplan Reconstruction Using Two-Level Queries", in *Conference on Computer Vision and Pattern Recognition*, 2023.

[43] J.-W. Su, K.-Y. Tung, et al., "SLIBO-Net: Floorplan Reconstruction via Slicing Box Representation with Local Geometry Regularization", in *Conference on Neural Information Processing Systems*, 2023.

[44] J. Chen, Y. Qian, and Y. Furukawa, "HEAT: Holistic Edge Attention Transformer for Structured Reconstruction", in *Conference on Computer Vision and Pattern Recognition*, IEEE/CVF, 2022.

[45] A. Gueze, M. Ospici, D. Rohmer, and M.-P. Cani, "Floor Plan Reconstruction from Sparse Views: Combining Graph Neural Network with Constrained Diffusion", in *International Conference on Computer Vision*, IEEE/CVF, 2023.

[46] A. Radford, J. W. Kim, C. Hallacy, et al., "Learning Transferable Visual Models From Natural Language Supervision", in *Proceedings of the International Conference on Machine Learning*, PMLR, 2021.

[47] F. L. Busch, T. Homberger, J. Ortega-Peimbert, Q. Yang, and O. Andersson, *One Map to Find Them All: Real-time Open-Vocabulary Mapping for Zero-shot Multi-Object Navigation*, 2024. arXiv: `2409.11764`.

[48] M. Luperto, M. M. Ferrara, G. Boracchi, and F. Amigoni, "Estimating Map Completeness in Robot Exploration", 2024. arXiv: `2406.13482`.

[49] C. Ho, S. Kim, B. Moon, A. Parandekar, N. Harutyunyan, C. Wang, K. Sycara, G. Best, and S. Scherer, "MapEx: Indoor Structure Exploration with Probabilistic Information Gain from Global Map Predictions", 2024. arXiv: `2409.15590`.

[50] A. Q. Li, F. Amigoni, and N. Basilico, "Searching for Optimal Off-Line Exploration Paths in Grid Environments for a Robot with Limited Visibility", in *Proceedings of the AAAI Conference on Artificial Intelligence*, 2012.

[51] P. E. Hart, N. J. Nilsson, and B. Raphael, "A Formal Basis for the Heuristic Determination of Minimum Cost Paths", *Transactions on Systems Science and Cybernetics*, vol. 4, no. 2, pp. 100–107, 1968.

[52] Z. Meng, H. Qin, Z. Chen, X. Chen, H. Sun, F. Lin, and M. H. Ang, "A Two-Stage Optimized Next-View Planning Framework for 3-D Unknown Environment Exploration, and Structural Reconstruction", *Robotics and Automation Letters*, vol. 2, no. 3, pp. 1680–1687, 2017.

[53] B. Zhou, Y. Zhang, X. Chen, and S. Shen, "FUEL: Fast UAV Exploration using Incremental Frontier Structure and Hierarchical Planning", *Robotics and Automation Letters*, 2021.

[54] S. M. LaValle, "Rapidly-Exploring Random Trees: A New Tool for Path Planning", Department of Computer Science, Iowa State University, Technical Report, 1998.

[55] S. Karaman and E. Frazzoli, "Sampling-based Algorithms for Optimal Motion Planning", *International Journal of Robotics Research*, vol. 30, no. 7, 2011.

[56] L. E. Kavraki, P. Svestka, J.-C. Latombe, and M. H. Overmars, "Probabilistic Roadmaps for Path Planning in High-Dimensional Configuration Spaces", *Transactions on Robotics and Automation*, vol. 12, no. 4, pp. 566–580, 1996.