# 1 Introduction

The first mathematical formulations of the problem of search appear at the end of the 17th century in the works by Huygens and Leibnitz [43] who required to find a trajectory of the searcher who pursues the target and tries to chase it. In the 20th century, this problem was considered both in its original formulation [61] and as a search game [27] in which the target also sees the searcher and tries to escape.

The other formulation of the problem was suggested in 1942 by Koopman [39] who considered distribution of search efforts over the domain and suggested his famous random search formula, which originated probabilistic search methods. In this formulation the goal is to detect the target or, in other words, to recognize location of the hidden target. Later Stone [55] suggested the theory of optimal search, which together with the Washburn pursuit theory [61] forms a basis for the development of search algorithms.

In this book, we consider the problem of search for static and moving targets by a single searcher and by a group of searchers. We assume that the searchers have imperfect information about locations of the targets and need to handle this uncertainty. The goal is to find a search policy such that the targets are detected in minimal time.



## 1.1 The Problem of Search Under Uncertainty

Modern studies of the methods of search for hidden objects follow two main directions. The first was originated in 1942 by the Koopman and the US Navy's Antisubmarine Warfare Operations Research Group with the aim of detecting German submarines in the Atlantic [39, 40]. As was later formulated by Frost and Stone [15, p. 13], the main question of these studies is

> how to most effectively employ limited resources when trying to find an object whose location is not precisely known. The goal is to deploy search assets to maximize the probability of locating the search object with the resources available. Sometimes this goal is stated in terms of minimizing the time to find the search object.

In the original scenario considered by Koopman, the searcher starts with a certain distribution of probabilities of the target's location in the domain and is required to distribute the search efforts over a domain, given these probabilities. It was assumed that the target is static and that the search efforts are infinitely divisible. After distribution of the search efforts, the points of the domain are observed, and the target is either detected or not. The goal is to find such a distribution of the search efforts that the probability of detecting the target is maximal.

To solve this problem, Koopman suggested his famous random search formula [39]

$$\varphi(x,t) = 1 - \exp(-\kappa(x,t)),$$

which defines probability $\varphi(x, t)$ of detecting the target in the point $x$ at time $t$ with respect to the search effort $\kappa(x, t)$ applied to the point $x$ at time $t$ and given that the target is located in the point $x$.

The Koopman random search formula states that the probability of detecting the target at a certain point is as great as the applied search effort, and perfect detection is resulted by application of infinite search efforts.

Starting with the Koopman research, in 1975 Stone [55] published the theory of optimal search and suggested an algorithm for creating optimal distribution of search efforts over the search domain. This algorithm considers the distribution of the probabilities of target's locations and iteratively defines the quantity of the search efforts that should be applied to each point of the domain. The Stone algorithm implements optimization using Lagrange multipliers and starting from the first point iteratively assigns the value of search efforts to the points of the domain proportionally to the logarithm of the location probabilities.

In 1980, Washburn [59, 60] developed the algorithm that defines optimal distribution of search efforts in the search for a moving target. This algorithm considers the probabilities of detecting and non-detecting the target and follows, respectively, forward and backward recursive equations for calculating these probabilities. Distribution of search efforts is defined by inverse Koopman formula such that it maximizes the probability of catching the target. As a result, given the probabilities of the target's movements, the algorithm creates steady-state distribution of the search efforts.

The second direction in the studies of search methods considers mobile searchers moving over the domain and seeking hidden static or moving targets. In contrast to the search games, in these studies it is assumed that the targets are not aware of the searchers' actions. Detection of the targets does not require catching the target and results from observation of the parts of the domain. Detection probability depends on the applied detection function that represents the used sensors.

For arbitrary function of detection probability in 1969, Ross [51] suggested a model that follows the direction of statistical decision processes and uses dynamic programming optimization. Later such models were formulated in terms of Markov Decision Processes [64] that allowed to consider imperfect detections with the errors of the first and the second type and to study the search for static and moving targets.

Solution for the simplest scenario of search for moving target was published in 1970 by Pollock [47], who considered a search for a target moving between two boxes and found an optimal policy for the searcher moves such that the target will be caught in minimal time. In later studies, the Pollock approach was extended for the domain with arbitrary number of points (for brief overview and appropriate references, see e.g. [35]).

In recent books, Stone et al. [56] and Washburn and Kress [63] considered both the indicated algorithms of probabilistic search and their extensions and suggested algorithms of search for moving targets [57], which combine the Koopman approach and the search by mobile agents. Along with that, a general solution to the search problem was not suggested yet, and the statement by Ross [52] that the searching for a moving target is an open problem remains valid.

## 1.2 Book Objectives and Main Contributions

The book addresses the search for static and mobile targets by the group of searchers, which are considered as autonomous agents – mobile robots equipped with certain sensors and governed by onboard controllers. It is assumed that in the group the agents can communicate with each other and share information about the targets detected.

In this book, we consider two types of algorithms:

- *reactive algorithms*, which govern the agents to plan their next actions following the information perceived by the sensors,
- *deep Q-learning algorithms*, which govern the agents to learn the patterns of the targets' motion and to plan their actions with respect to the learned patterns.

It is assumed that the sensed information is not always correct and that detection allows both false-negative and false-positive errors, which means that there is a nonzero probability that the agent does not recognize the existing target and that there is a nonzero probability that the agent recognizes nonexistent target.

The goal of the book is to present recently developed procedures for navigation of mobile agents searching for multiple targets in the presence of detection errors of the first and the second types. These procedures were developed with respect to different detection tasks and scenarios and are implemented in the form of constructive algorithms for detecting multiple static and mobile targets by the group of mobile agents acting under uncertainty.

The presented *reactive online algorithms* implement the strategies based on the "center of gravity" of the probability map and on the "center of view" of the used sensors and plan one-step or multi-steps motion using expected information gain as a decision-making criterion. The most effective reactive algorithm is an algorithm that implements dynamic Voronoi partitioning of the search domain and maximizes a one-step expected information gain in each Voronoi region.

In the presented *algorithms with deep Q-learning* abilities, it is assumed that the search agents are equipped with the strong enough computers processing prediction and target neural networks, which allow learning of the targets' activity. In the single-agent search, both networks are updated by the agent using the information obtained by its sensors. In the multi-agent search, the agents are allowed to communicate such that the next-in-line neighboring agent updates its prediction network with respect to the target network of the previous-in-line agent. Training of the networks can be conducted offline with an online correction or online starting with certain initial probability maps, which are corrected during the search.

The algorithms of both types are applicable in the search for multiple static and moving targets in the presence of the sensing errors of the first and the second types.

## 1.3 Structure of the Book

The book includes eight chapters, and the rest of the book is organized as follows. First, we form a background for further material and present three algorithms of search for static and moving targets under uncertainty, which illustrate the considered problem. Then, we formulate the problem and introduce the required terminology and basic concepts. The next chapter addresses reactive search procedures, and the final part deals with the algorithms with deep Q-learning.

More precisely, the contents of the chapters are the following.

Chapter 2, *Background*, briefly describes the considered search and detection problem and includes brief literature review. In addition, it includes three reactive algorithms of search: the first is a simple algorithm with information criterion, which resolves the uncertainty using diffusion process; the second is an algorithm of search for static targets with detection errors of the first and the second types; and the third is the algorithm of search in shadowed space.

Chapter 3, *Problem Formulation and Basic Search Procedure*, includes formalization of the considered problem of search and detection in the gridded domain, which is used in all scenarios throughout the book. Also, the chapter presents exact procedures for updating the probability maps and sensor fusion and specifies relations between probability maps of different levels, dynamics of sensors sensitivity, and communication between the agents. Finally, it presents the basic search procedure in the presence of false-positive and false-negative sensing errors.

Chapter 4, *Reactive Detection Algorithms*, presents the extensions of the formulated basic procedure and considers reactive algorithms of search. In these algorithms, the search agents are governed toward either the "center of gravity" of the probability map or the "center of view" of the sensors with one-step or multi-steps motion planning. Decision-making in both cases is based on the expected information gain. These algorithms are applied for the search by a single agent and by the groups of agents.

Chapter 5, *Search with Learning*, starts the consideration of search algorithms with the ability to learn the targets' motion patterns. The chapter presents a brief overview of supervised and unsupervised learning in the framework of search by mobile agents. This preliminary information is used in the next two chapters.

Chapter 6, *Search with Deep Q-learning: Single Agent*, presents the problem of search in the form of dynamic programming procedure. Based on this procedure, a deep reinforcement learning algorithm of search by a single agent is formulated. This algorithm implements deep Q-learning, which is used for maximizing cumulative gain of information about targets' locations and for minimizing the length of the searcher's trajectory. The algorithm terminates while a predefined detection probability is reached.

Chapter 7, *Search with Deep Q-learning: Multiple Agents*, presents an extension of the algorithm considered in the previous chapter to the search by multiple agents. To obtain such an extension, at first, a reactive algorithm based on the dynamic Voronoi partitioning of the search space is formulated. This algorithm plans the agents' paths such that they maximize the expected one-step look-ahead information in the agents' Voronoi regions. Then, the algorithm of search with deep Q-learning is formulated. In this algorithm, it is assumed that each search agent controls its prediction and training networks and that each agent updates its prediction network using the information stored in the target network of its previous-in-line neighboring agent.

Chapter 8, *Conclusions*, summarizes the discourse and presents further directions of the research.

All chapters that consider algorithms or search include numerical simulations and examples.

The book is self-contained and includes ready-to-use algorithms. Along with that, we consider it as an element of the array of publications, which fills a certain gap in the search theory but does not present complete search theory or even introduction to it.

To obtain additional information on the search theory, we recommend the report [7] and the report and the books by Stone and his colleagues [15, 55, 56]; the report [62] and the books [61, 63] by Washburn and his colleague; the book by Stone, Royset, and Washburn [57]; and our books [35, 36]. These books also include descriptions and references to the studies of the other approaches to search problems: bio-mimetic search, search in continuous space and time, and search based on path-planning algorithms.

The Python code of the main algorithms considered in the book can be downloaded from the page [20].