

# RoboSurgery

Matteo Nunziante\*

December 3, 2024

## **Abstract**

**Keywords:** keyword1, keyword2, keyword3

---

\*matteo.nunziante@studenti.units.it

# Contents

<b>1</b>	<b>Introduction</b>	<b>3</b>
<b>2</b>	<b>Related Work</b>	<b>3</b>
<b>3</b>	<b>Problem Statement</b>	<b>3</b>
<b>4</b>	<b>POMDP</b>	<b>3</b>
<b>5</b>	<b>Mathematical Formulation</b>	<b>4</b>
5.1	State Space . . . . .	4
5.2	Action Space . . . . .	4
5.3	Observation Space . . . . .	4
5.4	Bayesian estimation . . . . .	4
<b>6</b>	<b>Results</b>	<b>5</b>
6.1	MDP Solution . . . . .	5
6.2	QMDP . . . . .	5
6.3	Thompson Sampling . . . . .	5
6.4	Infotaxis . . . . .	6
6.5	Information Directed Sampling . . . . .	6
6.6	DQN . . . . .	6

# 1 Introduction

The goal of this project is to simulate a surgical robot for exploring lungs of patients.

# 2 Related Work

# 3 Problem Statement

# 4 POMDP

Formal description of Partially Observable Markov Decision Process as in [3]

Formally, a POMDP is a 7-tuple  $(S, A, T, R, \Omega, O, \gamma)$ , where:

$S$  is a set of states,

$A$  is a set of actions,

$T$  is a set of conditional transition probabilities between states,

$R : S \times A \rightarrow \mathbb{R}$  is the reward function.

$\Omega$  is a set of observations,

$O$  is a set of conditional observation probabilities,

$\gamma \in [0, 1)$  is the discount factor.

## 5 Mathematical Formulation

### 5.1 State Space

We are interested in the estimation of the deformation parameters of an object in a 2D space.

$$s = (\theta)$$

$\theta$  represents the deformation tensor.

### 5.2 Action Space

The action space is the set of all possible actions that the robot can take, Forwards, Backwards, Left, Right.

### 5.3 Observation Space

The observation space is the set of all possible observations that the robot can make, i.e. the presence or absence of an obstacle in its field of view.

### 5.4 Bayesian estimation

rendere il belief indipendente dalla posizione e orientazione del robot, dipenderò unicamente dalla deformazione

$$b(s) = b(\theta) * \delta(xyphi = real(xyphi))$$

$real(xyphi) = x^*$  from now on  $x = xyphi$

$$\begin{aligned} b'_{a,o}(x', \theta') &= \eta \cdot \delta(x' = x^* + a) \cdot p(o|x', \theta') \cdot \sum_{x, \theta} p(x', \theta'|x, \theta, a) b(x, \theta) \\ &= \eta \cdot \delta(x' = x^* + a) \cdot p(o|x', \theta') \cdot b(x' - a, \theta') \\ &= \eta \cdot \delta(x' = x^* + a) \cdot p(o|x', \theta') \cdot \delta(x' - a = x^*) b(\theta') \\ &= \eta \cdot \delta(x' = x^* + a) \cdot p(o|x', \theta') \cdot b(\theta') \\ &= \eta \cdot \delta(x' = x^* + a) \cdot b'_{a,o}(\theta') \end{aligned}$$

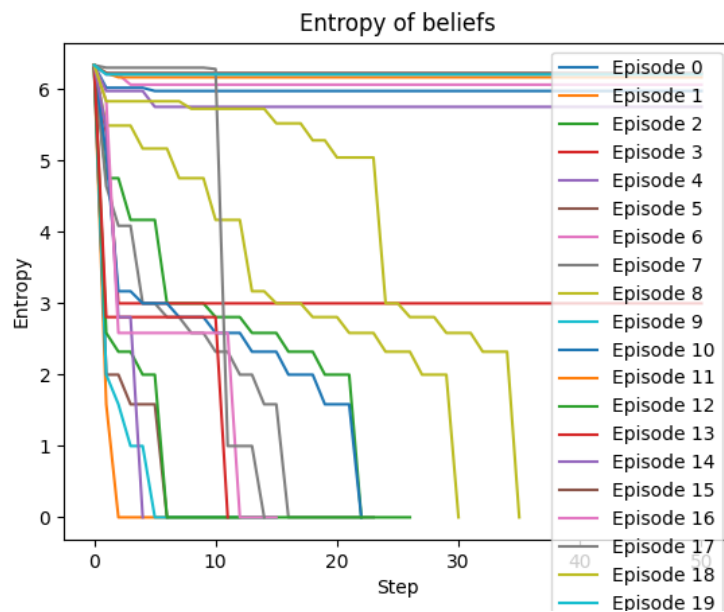
dunque a seconda della azione scelta i non zero saranno i 4 x raggiungibili con a

$$\begin{aligned} p(o|b, a) &= \sum_{x, \theta} p(o|x, \theta) b(x, \theta) \\ &= \sum_{\theta} p(o|x^*, \theta) b(\theta) \end{aligned}$$

## 6 Results

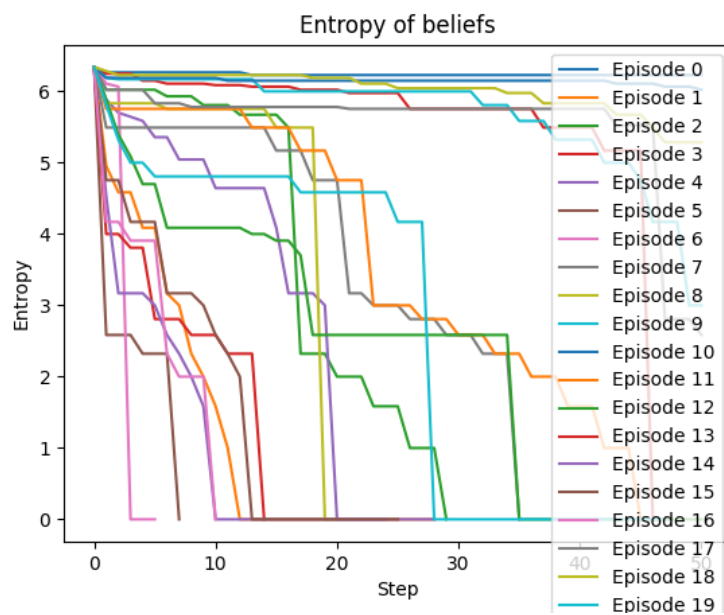
### 6.1 MDP Solution

### 6.2 QMDP



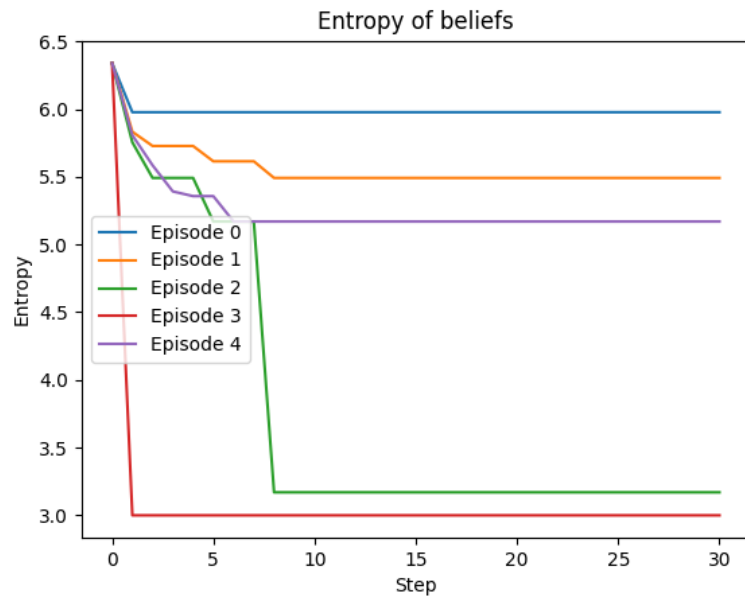
reaching target 13 out of 20 times.

### 6.3 Thompson Sampling



reaching target 14 out of 20 times.

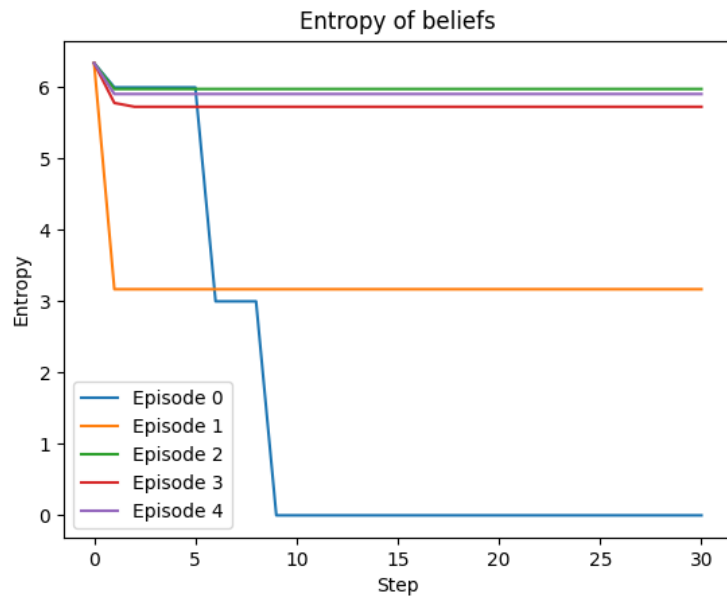
## 6.4 Infotaxis



reaching target 15 out of 20 times.

## 6.5 Information Directed Sampling

[2]



reaching target 0 out of 5 times

## 6.6 DQN

## References

- [1] Aurore Loisy and Christophe Eloy. Searching for a source without gradients: how good is infotaxis and how to beat it. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 478(2262), June 2022.
- [2] Daniel Russo and Benjamin Van Roy. Learning to optimize via information-directed sampling, 2017.
- [3] Matthijs T. J. Spaan. *Partially Observable Markov Decision Processes*, pages 387–414. Springer Berlin Heidelberg, Berlin, Heidelberg, 2012.