

RESEARCH ARTICLE

# Robust Satisficing Decision Making for Unmanned Aerial Vehicle Complex Missions under Severe Uncertainty

Xiaoting Ji, Yifeng Niu\*, Lincheng Shen

College of Mechatronics and Automation, National University of Defense Technology, Changsha, Hunan, China

\* [niuyifeng@nudt.edu.cn](mailto:niuyifeng@nudt.edu.cn)



**OPEN ACCESS**

**Citation:** Ji X, Niu Y, Shen L (2016) Robust Satisficing Decision Making for Unmanned Aerial Vehicle Complex Missions under Severe Uncertainty. PLoS ONE 11(11): e0166448. doi:10.1371/journal.pone.0166448

**Editor:** Yong Deng, Southwest University, CHINA

**Received:** April 7, 2016

**Accepted:** October 28, 2016

**Published:** November 11, 2016

**Copyright:** © 2016 Ji et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All relevant data are within the paper and its Supporting Information files.

**Funding:** This work was supported by National Natural Science Foundation of China, 61403410, <http://www.nsf.gov.cn/>, YN. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** The authors have declared that no competing interests exist.

## Abstract

This paper presents a robust satisficing decision-making method for Unmanned Aerial Vehicles (UAVs) executing complex missions in an uncertain environment. Motivated by the info-gap decision theory, we formulate this problem as a novel robust satisficing optimization problem, of which the objective is to maximize the robustness while satisfying some desired mission requirements. Specifically, a new info-gap based Markov Decision Process (IMDP) is constructed to abstract the uncertain UAV system and specify the complex mission requirements with the Linear Temporal Logic (LTL). A robust satisficing policy is obtained to maximize the robustness to the uncertain IMDP while ensuring a desired probability of satisfying the LTL specifications. To this end, we propose a two-stage robust satisficing solution strategy which consists of the construction of a product IMDP and the generation of a robust satisficing policy. In the first stage, a product IMDP is constructed by combining the IMDP with an automaton representing the LTL specifications. In the second, an algorithm based on robust dynamic programming is proposed to generate a robust satisficing policy, while an associated robustness evaluation algorithm is presented to evaluate the robustness. Finally, through Monte Carlo simulation, the effectiveness of our algorithms is demonstrated on an UAV search mission under severe uncertainty so that the resulting policy can maximize the robustness while reaching the desired performance level. Furthermore, by comparing the proposed method with other robust decision-making methods, it can be concluded that our policy can tolerate higher uncertainty so that the desired performance level can be guaranteed, which indicates that the proposed method is much more effective in real applications.

## 1 Introduction

Over the past decades, Unmanned Aerial Vehicles (UAVs) have been extensively employed in many civil and military applications, such as search and rescue in the hazardous environment [1, 2], environment surveillance [3], 3D terrain reconstruction [4], climate research [5], and ground reconnaissance [6]. Various types of missions require high autonomy for UAVs to

make proper decisions in complex environments. These complex missions may specify constraints on order, relative time, goal, safety, etc. Therefore, high-level mission specifications using temporal logics have been employed to improve the expressiveness of complex missions in recent years, such as the Linear Temporal Logic (LTL) [7–9]. The LTL provides a formal high-level framework to specify the complex mission with a natural encoding of Boolean and temporal operators, and atomic propositions, which will help enhance the interaction with UAVs and realize a better understanding of the behaviors of UAVs. For instance, a complex mission performed by an UAV can be described as “Take off from home A, search regions B, C, and D in a given order while always avoiding the unsafe regions F, and finally return to A”. Moreover, most missions for UAVs take place in complex and uncertain environments, which will result in actuation or sensing uncertainties for the UAVs. How to make proper decisions in face of uncertainties is an important aspect for UAVs decision making. Motivated by this situation, some researchers have studied the Markov Decision Process (MDP), which provides a general mathematical framework for sequential decision making under uncertainty [10]. MDPs use transition probabilities to model behavior uncertainties, which are caused by actuation errors or environmental disturbances. Considering the complex missions for the UAVs in a complex and uncertain environment, the synthesizing control method by combining the MDP with LTL have been studied. The decision making problem for the UAV executing a complex mission can be translated to generate a policy for the MDP in order to satisfy the LTL specifications. Many synthesizing control algorithms [11–13] based on the model checking theory [14, 15] have been provided in order to find an optimal control policy for the MDP with maximal probability of satisfying the LTL specifications.

However, when a real UAV system is abstracted to the MDP, it may lead to modeling uncertainties. And it is often prohibitively costly or even infeasible to obtain accurate transition probabilities in practice. Thus the estimated transition probabilities may deviate far from the true value due to errors in the abstraction process. Based on the existing works, it turns out that the optimal policy is often quite sensitive to even minor errors in the transition probabilities [16]. Therefore, the assumption that the transition probabilities are exact, when in fact they are uncertain, can lead to suboptimal decisions, or even degradation of system performance [17]. For synthesizing the MDP with LTL specifications, the uncertain transition probabilities will lead to a low probability of satisfying the LTL specifications, which means it will increase risks such as UAVs destruction and mission failure. So it is very important that the UAV must make robust decisions with respect to the uncertain transition probabilities of the MDP. Many researchers have devoted efforts to modeling the uncertain transition probability and mitigating its effect on the resulting policy in order to satisfy the mission requirements. The study of the MDP with uncertain transition probabilities, which is called an uncertain MDP, can date back to the 1970s [18–20]. There are many uncertainty models to describe the uncertain transition probabilities. The most common approach is to assume that the uncertain transition probabilities lie in an interval [13, 21]. The interval model is motivated by statistical estimates of confidence intervals on the individual components of the transition probability. Other statistic uncertainty models, which have been explicitly described in [22], include the likelihood model, the Bayesian model, and the entropy model. Relevant algorithms have been developed to handle uncertain MDPs, which mainly employ the game theory, the queuing theory, or dynamic programming. The most commonly used method is the min-max robust decision-making method [22], where the principle of optimality criterion is to maximize the worst-case expected total utility. Based on this optimality criterion, robust value iteration and robust policy iteration [23, 24] algorithms are proposed to obtain a min-max robust policy, and they both depend on robust dynamic programming, which may address the issue of designing an approximation method

with an appropriate robustness to extend the power of the Bellman Equation. Iyengar [25] and Nilim and El Ghaoui [22] suggest to find a policy to guarantee the highest expected total utility at a given confidence level. To this end, a policy that maximizes the worst-case objective is determined. Bertuccelli et al [26] develop a robust adaptive MDP to optimize the UAV decisions. Although a great deal of researches have been done, robust decision making for uncertain MDPs with LTL specifications is rarely considered. To our knowledge, there are only a few studies on this problem. Wolff et al [27] propose a robust version of dynamic programming to work out a min-max robust policy, which maximizes the worst-case probability of satisfying LTL specifications. The previous robust methods are disadvantageous in several aspects. First, the uncertainty models require adequate knowledge of the uncertainty range (e.g., the interval model) or probabilistic information (e.g., Bayesian and likelihood models) in advance. It cannot handle the situation where the uncertainty is fuzzy, non-probabilistic or unknown, which is called severe uncertainty [28]. Since severe uncertainty often occurs in uncertain and dynamic environments [29], UAVs have to make robust decisions in order to reduce the influence of severe uncertainty. Second, the min-max robust decision-making method may lead to an overly conservative resulting policy, which indicates the policy will be sub-optimal in any case but the worst case. Furthermore, determining the worst case would become infeasible when the uncertainty is severe.

Motivated by the previous two limitations, another robust decision theory, namely the info-gap decision theory [28], is proposed to solve the robust decision-making problem from another perspective, which has been applied in a large spectrum of fields, including engineering, biology, and project management. It is a conceptual framework that can support making decisions under severe uncertainty and that has substantive implications for the formulation, evaluation and selection of desired goals and of the means to attain them. First, the uncertainty model is non-probabilistic, defined by a family of nested sets. That is, it does not require one to specify either a probability distribution or bounds on the uncertainty set. Second, the info-gap decision theory derives a robust satisficing policy by maximizing the robustness (i.e., the immunity to uncertainty) while satisfying a desired performance level, instead of seeking for the optimal utility. The term 'satisficing', a combination of satisfy and suffice [30], is introduced by Simon in 1956 [31]. Satisficing is a decision-making strategy aimed at a satisfactory or adequate result, rather than the optimal solution. For our problem, when the UAV executes complex civil or military missions, human operators often specify some critical mission criteria instead of the optimal ones, because the optimality is often difficult to attain or costly in terms of pay in real applications. Moreover, the robustness is a reliable principle often used in complex decision-making problems involving severe uncertainty [32, 33] because it provides the maximum reduction of unknown risks. Therefore, we will propose a robust satisficing decision-making method based on the info-gap decision theory to solve the robust decision-making problem for the UAV executing complex missions under uncertainty, i.e., the synthesizing control of uncertain MDPs with LTL specifications. The goal is to generate a robust satisficing policy that can maximize the robustness while ensuring the desired probability level of satisfying the LTL specifications. There are several works using the info-gap decision theory to obtain a robust satisficing policy for their research areas, such as neural network [34], multi-agent search [35], and path planning [36]. S. Gareth et al [34] use an info-gap model to quantify the network response to uncertainty in the input data in order to evaluate the reliability of the neural network. Itay et al [35] propose a robust satisficing approach based on the info-gap decision theory to solve the spatial search-planning problem with imprecise probabilistic data. Mascareñas et al [36] develop a path planner anchored in the info-gap decision theory to generate non-deterministic paths that satisfy predetermined performance requirements in the face of the uncertain actions of the hostile elements. However, all the previous works do not

consider the system model as a MDP and the complex mission requirements, even with regard to the synthesizing control problem of uncertain MDPs with LTL specifications. Therefore, utilizing the info-gap decision theory to solve robust UAV decision making for complex missions under severe uncertainty is still an open problem.

Our work is carried out based on the combination of the MDP and the info-gap decision theory. The LTL is utilized to describe the complex mission requirements of UAVs, which is introduced as constraints of the optimization problem. The main contribution of this paper is as follows. A robust satisficing decision-making method based on the info-gap decision theory is proposed for the robust UAVs decision-making problem with complex mission requirements under severe uncertainty. First, motivated by the info-gap decision theory, we propose a novel optimization problem for the robust UAVs decision-making problem, which can maximize the robustness while ensuring the desired mission requirements. Specifically, the uncertain UAV system is modeled as a new info-gap based MDP, and a robustness function is formulated to evaluate the robustness with the LTL formula specifying the mission specifications, so a robust satisficing policy is obtained to achieve the maximal robustness. To our knowledge, this is the first work that extends the robust satisficing concept into UAV robust decision making under severe uncertainty, with consideration given to critical mission specifications. Second, a two-stage robust satisficing solution strategy is proposed to solve the previous problem, which consists of the construction of a product IMDP and the generation of a robust satisficing policy. In the first stage, the product IMDP is constructed by creating the Cartesian product of the IMDP and DRA converted from the LTL formula, which is used to compute the probability of satisfying the LTL formula. In the second stage: i) the monotonic relationship between the uncertainty level and the worst-case probability of satisfying the LTL formula is provided and proven, so that when the uncertainty level continuously increases, the worst-case probability will reach the critical satisfying condition of the desired performance level; ii) based on the previous monotonic relationship, the robust satisficing optimality theorem is proved, which can help improve the policy towards the direction with a higher robustness; and iii) an algorithm based on robust dynamic programming is proposed to generate a robust satisficing policy, while an associated robustness evaluation algorithm is presented to evaluate the robustness. Finally, through Monte Carlo simulation, the effectiveness of our algorithm is demonstrated on an UAV search mission under severe uncertainty so that the resulting policy can maximize the robustness while satisfying the desired performance level. Further, by comparing the proposed method with the min-max robust decision-making method and the robust decision making [37] (another robust optimization method for handling the severe uncertainty), it can be concluded that our policy can tolerate higher uncertainty so that the desired performance level can be guaranteed, which indicates that the proposed method is much more effective in real applications of the UAVs decision making problem.

The remainder of this paper is organized as follows. Some preliminary definitions are presented in Section 2. The problem is formulated based on the info-gap decision theory and the solution method is outlined in Section 3. In Section 4, a robust satisficing policy generation algorithm based on robust dynamic programming is proposed as well as a robustness evaluation algorithm. In Section 5, our algorithms are illustrated by an example of the UAV search mission, and the results are analyzed. The conclusion and future work are presented in Section 6.

## 2 Preliminaries

In this section, we will present some preliminary definitions that will be used in this paper.

## 2.1 LTL Specifications

The LTL formula is used to specify the complex mission with temporal constraints.

**Definition 1 (LTL [14]):** An LTL formula  $\phi$  can be defined recursively by a set of atomic propositions  $AP$  and a set of unary and binary operators:

$$\phi ::= p \mid \neg p \mid \phi \vee \phi \mid \phi \wedge \phi \mid X \phi \mid \phi U \phi \mid F \phi \mid G \phi$$

where  $p \in AP$  is an atomic proposition;  $\neg$  (negation),  $\vee$  (disjunction), and  $\wedge$  (conjunction) are the standard Boolean operators;  $X$  (next),  $U$  (until),  $F$  (eventually), and  $G$  (always) are the temporal operators.

The semantics of LTL formulas are defined over infinite words in  $2^{AP}$ . Given an infinite word  $\sigma = \tau_0\tau_1 \dots \tau_i \dots$ , if the LTL formula  $\phi$  is true at the first position, we say the word satisfies the LTL formula  $\phi$ , denoted as  $\sigma \models \phi$ .  $X\phi$  means that  $\phi$  becomes true at the next position of the word;  $G\phi$  means that  $\phi$  is true at all positions of the word;  $F\phi$  means that  $\phi$  eventually becomes true in the word;  $\phi_1 U \phi_2$  means that  $\phi_1$  holds at each position in the word until  $\phi_2$  is true.

In quantitative probabilistic verification, an LTL formula needs to be translated into a deterministic Rabin automaton (DRA) by the PRISM [38], a leading probabilistic model checker.

**Definition 2 (DRA [14]):** A DRA is a tuple  $A_\phi = \{Q, q_0, \Sigma, \delta, Acc\}$ , where  $Q$  is a finite set of states,  $q_0 \in Q$  is the initial state,  $\Sigma = 2^{AP}$  is an input alphabet,  $\delta: Q \times \Sigma \rightarrow Q$  is the transition function, and  $Acc \subseteq 2^Q \times 2^Q$  is a set of accepting state pairs.

Let  $\omega = \omega_0\omega_1 \dots$  be a string over  $\Sigma$ . A run  $\omega$  denotes an infinite sequence  $q_0q_1 \dots$  of states in  $A_\phi$  such that  $q_{i+1} = \delta(q_i, \omega_i)$  for  $i \geq 0$ . The run is accepted by a DRA, if for an accepting pair  $(J, K) \in Acc$ , the set of states  $J$  is finitely often visited and the set of states  $K$  is infinitely often visited.

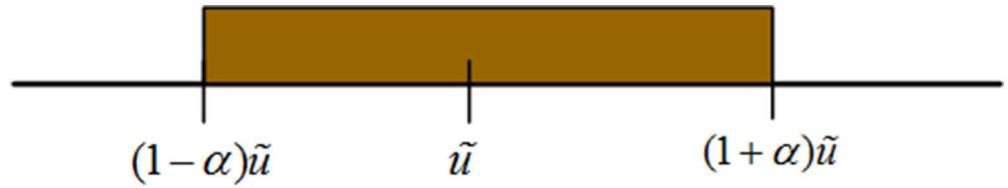
## 2.2 System Model

We consider the UAV system with noisy actuation of which the dynamics are described by a stochastic differential equation [13]. It is assumed that the evolution of the stochastic system satisfies the Markov property, which can be abstracted to an MDP.

**Definition 3 (MDP [12]):** A (labeled) MDP is defined as a tuple  $\langle S, A, P, R, s_0, AP, L \rangle$ , where  $S$  is a finite set of states;  $A$  is a finite set of actions ( $A(s) \subseteq A$  denotes the actions available at state  $s \in S$ );  $P: S \times A \times S \rightarrow [0, 1]$  is the transition probability function, such that for all  $s \in S$ ,  $\sum_{s' \in S} P(s, a, s') = 1$  if  $a \in A(s)$ , and  $P(s, a, s') = 0$  if  $a \notin A(s)$ ;  $R: S \times A \rightarrow \mathbb{R}^+$  is the reward function;  $s_0 \in S$  is the initial state;  $AP$  is a finite set of atomic propositions; and  $L: S \rightarrow 2^{AP}$  is a labeling function that establishes which atomic propositions are true in the given state  $s \in S$ , i.e.,  $L$  relates discrete states with the proposition regions.

We use  $P_{ij}^a$  as shorthand for the transition probability from state  $i$  to state  $j$  when using action  $a$ .  $P^a: S \times S \rightarrow [0, 1]$  represents a transition matrix, where the  $(i, j)$ -th entry of  $P^a$  is  $P_{ij}^a$ . When the transition probabilities of the MDP are uncertain, it is only known that the corresponding transition matrix  $P^a$  for each action  $a$  lies in some given subset  $\mathcal{P}^a$ .

**Definition 4 (Uncertain MDP [22]):** The uncertain MDP is defined as  $uM = \langle S, A, \mathcal{P}, R, s_0, AP, L \rangle$ , where  $\mathcal{P} = \mathcal{P}^1 \times \dots \times \mathcal{P}^{|A|}$  is the uncertainty set for the transition probabilities. For all  $a \in A$ , it is assumed that the sets  $\mathcal{P}^a$  satisfy the rectangular uncertainty property, i.e.,  $\mathcal{P}^a = \mathcal{P}_1^a \times \dots \times \mathcal{P}_n^a$ .



**Fig 1. The space of possible values of uncertain variables.**

doi:10.1371/journal.pone.0166448.g001

### 2.3 Info-gap Decision Theory

The info-gap decision theory consists of three components: an info-gap uncertainty model, a robustness function, and a robust satisficing policy.

**Definition 5 (Info-gap uncertainty model [28]):** Assuming the best estimation of an uncertain parameter  $u$  is  $\tilde{u}$ , and the relative errors between these two values are unknown, the info-gap uncertainty model can be represented as a family of nested sets:

$$U(\alpha, \tilde{u}) = \{u : |u - \tilde{u}| \leq \alpha \tilde{u}, \alpha \geq 0\} \tag{1}$$

where  $\alpha$  is the unknown fractional deviation from the estimated value, i.e. the uncertainty level. In this paper, it is assumed that  $\tilde{u}$  is the estimated transition probability of the MDP, which will be explicitly described in the next subsection.

The set  $U(\alpha, \tilde{u})$  contains all parameters  $u$  of which fractional deviation from  $\tilde{u}$  is no greater than  $\alpha$ , as shown in Fig 1.

The info-gap uncertainty model obeys two axioms:

- Contraction:  $U(0, \tilde{u}) = \{\tilde{u}\}$ .
- Nesting:  $\alpha_1 \leq \alpha_2 \Rightarrow U(\alpha_1, \tilde{u}) \subseteq U(\alpha_2, \tilde{u})$ .

From the contraction property, we can see that  $\tilde{u}$  is the only value if there is no uncertainty. And the second property states that the higher the uncertainty level is, the more inclusive the info-gap uncertainty model will be.

The robustness function of the info-gap decision theory measures the highest uncertainty level for which a given policy will satisfy the performance requirements, thus allowing maximal lack in the knowledge of a priori information.

**Definition 6 (Robustness function [28]):** In the info-gap theory, the robustness function of a given policy  $\pi$  is defined as the highest level of uncertainty that can be tolerated, for which the given policy will satisfy a desired performance level  $r_c$

$$\hat{\alpha}(\pi, r_c) = \max \left\{ \alpha : \min_{u \in U(\alpha, \tilde{u})} R(\pi, u) \geq r_c \right\} \tag{2}$$

where  $U(\alpha, \tilde{u})$  is the info-gap uncertainty model, and  $R(\pi, u)$  is the performance evaluation function for policy  $\pi$  and uncertain parameter  $u$ .

**Definition 7 (Robust satisficing policy [28]):** A robust satisficing policy is defined as the policy that maximizes the robustness function Eq (2) while satisfying the desired performance level  $r_c$

$$\pi^*(r_c) = \arg \max_{\pi \in \Pi} \hat{\alpha}(\pi, r_c) \tag{3}$$

where  $\Pi$  represents the decision space that consists of a set of possible policies.

The robust satisficing policy  $\pi^*(r_c)$  maximizes  $\hat{\alpha}(\pi, r_c)$  conditional on the desired performance level  $r_c$ , and makes the condition  $R(\pi^*(r_c), u) \geq r_c$  guaranteed for any  $u \in U(\alpha, \tilde{u})$ .

### 3 Info-gap Based Robust Satisficing Decision-Making Problem

#### 3.1 Info-gap based MDP

In this paper, we use the info-gap uncertainty model to represent the uncertain transition probabilities of the MDP. Let  $P$  be the unknown true transition probability,  $\tilde{P}$  be the estimated transition probability of the MDP, and  $P^a$  and  $\tilde{P}^a$  be the transition matrices of  $P$  and  $\tilde{P}$  for taking action  $a$  respectively. The info-gap uncertainty model of the transition matrix for taking action  $a$  is defined as follows

$$U^a(\alpha, \tilde{P}) = \{P^a : |p_i - \tilde{p}_i| \leq \alpha \tilde{p}_i, 0 \leq \alpha \leq 1, p_i \mathbf{1} = 1, \tilde{p}_i \mathbf{1} = 1, p_i \geq 0, \tilde{p}_i \geq 0\} \forall i \quad (4)$$

where  $p_i$  and  $\tilde{p}_i$  represent the  $i$ th row of  $P^a$  and  $\tilde{P}^a$ , and  $\alpha$  is the unknown uncertainty level. The uncertainty set of  $p_i$  can be expressed as an interval  $[\underline{p}, \bar{p}]$ , where  $\underline{p} = (1 - \alpha)\tilde{p}$ ,  $\bar{p} = (1 + \alpha)\tilde{p}$ , and  $\underline{p}, \bar{p} \geq 0$ . Since  $\alpha$  is unknown, the interval model is not fixed. The range of uncertainty expands as  $\alpha$  increases.

**Definition 8 (IMDP):** The info-gap based (labeled) MDP (IMDP) is defined as a tuple  $IM = \langle S, A, U(\alpha, \tilde{P}), R, s_0, AP, L \rangle$  by replacing  $\mathcal{P}$  in the uncertain MDP with  $U$ , where  $U(\alpha, \tilde{P}) = U^1(\alpha, \tilde{P}^1) \times \dots \times U^{|A|}(\alpha, \tilde{P}^{|A|})$  are the sets of all possible transition matrices, as defined in Eq (4).

A control policy for the IMDP is defined as a sequence  $\pi = \{\mu_0, \mu_1, \dots\}$ , where  $\mu_i: S \rightarrow A$  is a control function such that  $\mu(s) \in A(s)$  for all  $s \in S$ . If  $\pi = \{\mu, \mu, \dots\}$ , the control policy is called a stationary policy.

#### 3.2 Problem Formulation

Considering both the robustness of an IMDP model and the probability of satisfying the LTL specifications, we propose an info-gap based robust satisficing decision-making method. The objective is to maximize the robustness to uncertainty in the IMDP model while guaranteeing the desired performance level satisfied. In this paper, the performance level is defined as the desired mission success rate of the UAV, i.e., the desired probability level of satisfying the LTL specifications.

**Definition 9 (LSP and DLSP):** The probability of satisfying the LTL specification  $\phi$  (LSP) by the IMDP  $IM$  under the control policy  $\pi$  is defined as  $Pr^\pi(s_0 \models \pi)$ . And the desired probability level of satisfying the LTL specification (DLSP) is set as a constant value within  $(0, 1)$  by the UAV operator.

A control policy  $\pi$  of the IMDP  $IM$  can produce a path  $r_{IM}^\pi = s_0 s_1 \dots s_i \dots$  over  $IM$ , which will further generate a corresponding infinite word  $\sigma = \tau_0 \tau_1 \dots \tau_i \dots$  over the atomic propositions of the LTL such that  $\tau_i = L(s_i)$ . Thus, the probability of satisfying the LTL formula  $\phi$  (LSP) for a path produced by policy  $\phi$  over  $IM$  can be represented as  $Pr^\pi(s_0 \models \pi) = Pr\{L(r_{IM}^\pi) \models \phi\}$ , which is measurable [14].

Formally, the robust satisficing decision-making problem can be formulated as follows:

**Problem 1:** Let  $IM = \langle S, A, U(\alpha, \tilde{P}), R, \gamma, s_0, L, AP \rangle$  be an IMDP, and  $\phi$  be an LTL formula over  $AP$ . The objective is to generate a robust satisficing policy  $\pi^*(p_c)$  that maximizes the

robustness while guaranteeing the DLSP  $p_c$  satisfied

$$\pi^*(p_c) = \arg \max_{\pi \in \Pi} \alpha \tag{5}$$

$$s.t. \min_{P \in U(\alpha, \tilde{P})} Pr^\pi(s_0 \models \phi) \geq p_c \tag{6}$$

where  $\Pi$  represents the decision space consisting of a set of possible policies, and  $Pr^\pi(s_0 \models \phi)$  is defined as the LSP of the LTL formula  $\phi$  by IM under the control policy  $\pi$  from an initial state  $s_0$ .

For a DLSP  $p_c$ , the robustness of a policy  $\pi$  can be defined as

$$\hat{\alpha}(\pi, p_c) = \max \alpha \tag{7}$$

$$s.t. \min_{P \in U(\alpha, \tilde{P})} Pr^\pi(s_0 \models \phi) \geq p_c \tag{8}$$

**Remark 1:** For a set of feasible policies, the preference can be determined by the robustness of each policy. The higher the robustness is, the higher the preference will be.

### 3.3 Robust Satisficing Decision-Making Framework

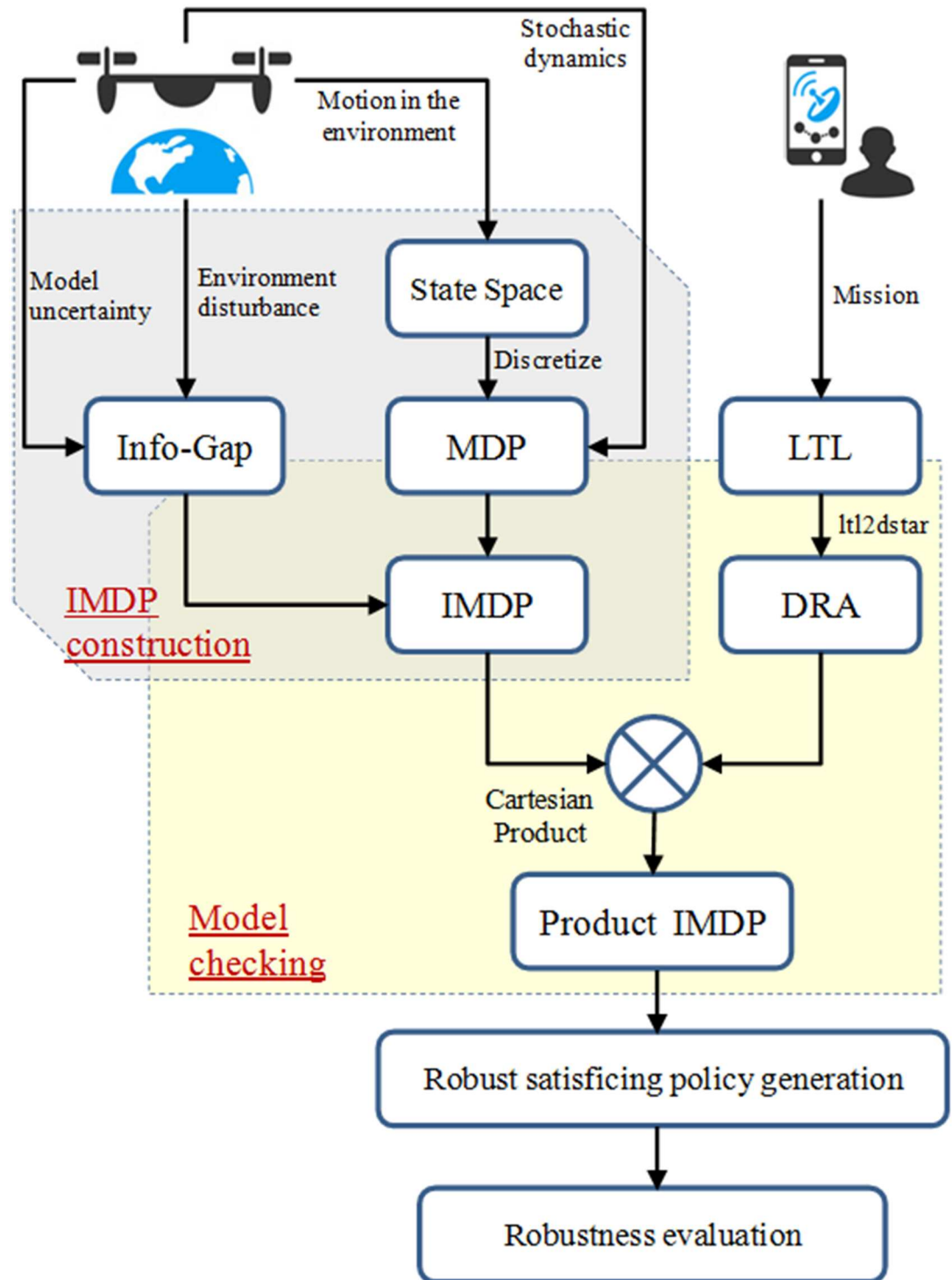
In this subsection, we will present a robust satisficing decision-making framework for Problem 1, which is a computational framework that produces a policy that maximizes the robustness to uncertainty while guaranteeing the DLSP satisfied, as shown in Fig 2. This framework consists of two main parts, the construction of the product IMDP and the generation of a robust satisficing policy, which will be described in detail in the following.

**3.3.1 Construction of the Product IMDP.** According to the model checking theory [14], the LSP for the IMDP is equivalent to the probability of reaching an accepting maximal end component (AMEC) of the product IMDP for a given policy and a given transition probability. By constructing the product IMDP, the LSP can be computed. Therefore, we start by constructing the product IMDP, which is used for computing the LSP in order to determine whether the DLSP in Problem 1 is satisfied or not. First, the LTL formula should be converted into a DRA as defined in Def. 2. The product IMDP is constructed by creating the Cartesian product of the IMDP and DRA.

**Definition 10 (Product IMDP):** Let  $IM = \langle S, A, U(\alpha, \tilde{P}), R, s_0, AP, L \rangle$  be an IMDP. The LTL formula is converted into a DRA as  $A_\phi = \{Q, q_0, \Sigma, \delta, Acc\}$ . The product IMDP is defined by  $P = \langle S_p, A_p, U_p(\alpha, \tilde{P}), R_p, s_{p0}, L_p, Acc_p \rangle$ , where

- $S_p = S \times Q$ .
- $A_p(s, q) = A(s)$ .
- For  $P_p \in U_p(\alpha, \tilde{P})$  and  $P \in U(\alpha, \tilde{P})$ ,  $P_p((s, q), a, (s', q')) = P(s, a, s')$  if  $q' = \delta(q, L(s))$ ; otherwise, 0.
- $R_p((s, q), a) = R(s, a)$ .
- $s_{p0} = (s_0, q_0)$ .
- $L_p(s, q) = q$ .
- $Acc_p = \{(J_1^P, K_1^P), (J_2^P, K_2^P), \dots\}$ . For  $(L_i, J_i) \in Acc$ , state  $(s, q) \in J_i^P$  if  $q \in J_i$ , and state  $(s, q) \in L_i^P$  if  $q \in L_i$ .





**Fig 2. Robust satisficing decision-making framework.**

doi:10.1371/journal.pone.0166448.g002

The policy  $\pi^p = \{\mu_0^p, \mu_1^p, \dots\}$  on the product IMDP is denoted as  $\mu_i^p : S \times Q \rightarrow A$ . There is a one-to-one correspondence between the paths on the IMDP and the product IMDP, which induces a one-to-one correspondence between the policies on the IMDP and the product IMDP. Therefore, given a policy  $\pi^p = \{\mu_0^p, \mu_1^p, \dots\}$  on the product IMDP, one can induce a policy  $\pi = \{\mu_0, \mu_1, \dots\}$  on the IMDP by setting  $\mu_i^p(s, q) = \mu_i(s)$  for  $i = 0, 1, \dots$ .

With the product IMDP, the detailed procedure of obtaining the AMECs is outlined in [14].

**Definition 11 (AMEC [14]):** The accepting maximal end component is defined as  $(\bar{S}_p, \bar{A}_p)$ , consisting of a set of states  $\bar{S}_p \subseteq S_p$  and a function  $\bar{A}_p(s_p) \subseteq A_p(s_p)$ , which implies that by taking actions enabled by  $\bar{A}_p$ , all states in  $\bar{S}_p$  can reach every other state in  $\bar{S}_p$  and cannot reach any state outside of  $\bar{S}_p$ .

Once an AMEC is reached, all states in  $\bar{S}_p$  are infinitely often reached with probability 1, by taking all actions in  $\bar{A}_p$ . The LSP is the maximum probability of reaching any states in  $\bar{S}_p$  from an initial state  $s_{p0} \in S_p$ . We can find the set of states that can never reach  $\bar{S}_p$  under any policy via the graph theory, denoted as  $B_0$ . The set of the rest states is  $B_p = S_p / (\bar{S}_p \cup B_0)$ . According to the model checking theory, the LSP from the initial state  $s_p$  can be determined as: 1 if  $s_p \in \bar{S}_p$ , or 0 if  $s_p \in B_0$ . For  $s_p \in B_p$ , the LSP can be obtained through linear or dynamic programming [14] [11] if there is no uncertainty. However, in this paper the system is abstracted into the IMDP model, and the transition probability lies in an info-gap uncertainty model, which requires the robust satisficing solution method. In the next subsection, the solution scheme will be presented for obtaining a robust satisficing policy for Problem 1.

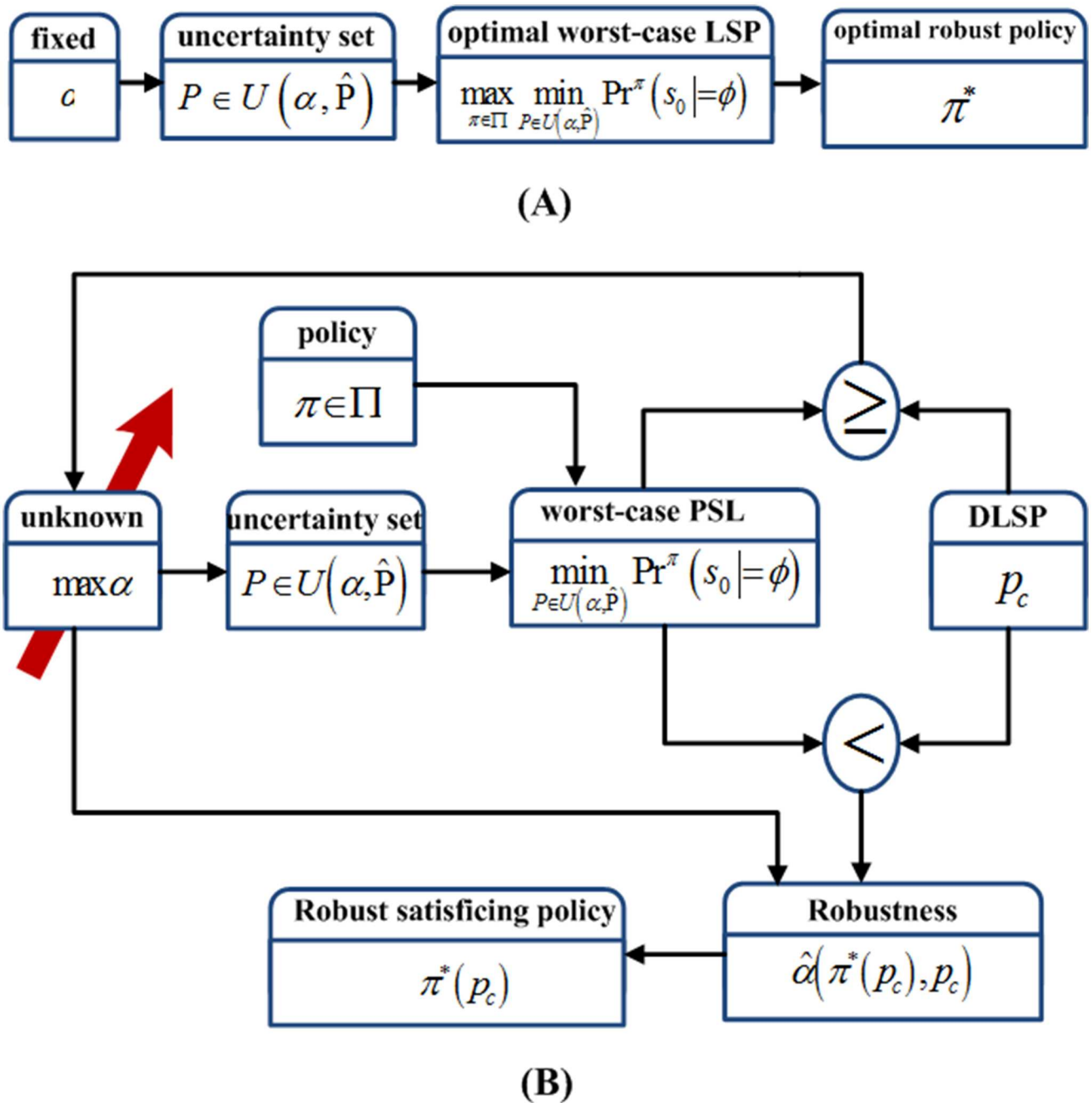
**3.3.2 Solution Scheme.** It can be seen that Problem 1 for finding the robust satisficing policy is essentially a complex optimization problem. Both the policy  $\pi$  and the uncertainty level  $\alpha$  will affect the value of the worst-case LSP, so they will affect determining whether the DLSP is satisfied or not. In order to clarify the solution procedure of generating a robust satisficing policy, it is compared with that of the min-max decision-making method. The solution schemes are shown in Fig 3.

Considering the IMDP model and the LTL specification, the min-max robust decision-making method is to create an optimal robust policy to maximize the worst-case LSP at a specified uncertainty level  $\bar{\alpha}$ . It can be regarded as a game between the environment and the decision maker, and the optimal robust policy is

$$\pi^* = \arg \max_{\pi \in \Pi} \min_{P \in U(\bar{\alpha}, \bar{P})} Pr^\pi(s_0 \models \phi) \tag{9}$$

This min-max optimization problem written as Eq (9) can be solved to find an optimal robust policy through robust dynamic programming. The solution scheme can be seen in Fig 3(A). However, it is only feasible when the uncertain transition probability set is fixed, i.e., the uncertainty level  $\bar{\alpha}$  is known in advance.

The solution scheme for Problem 1, i.e., the robust satisficing decision-making method, is shown in Fig 3(B). In order to simplify the computation, the uncertainty level within [0,1] will be divided into N uniform divisions. We start by choosing the uncertainty level  $\alpha_t$  as zero. For each specified uncertainty level  $\alpha_t$ , a policy with the highest worst-case LSP will be found, and the highest worst-case LSP will be compared with the DLSP. If the highest worst-case LSP is higher than or equal to the DLSP, the value of the uncertainty level will be increased by  $\alpha_{t+1} = \alpha_t + 1/N$ , and the above procedures will be repeated; otherwise, the last policy with the highest worst-case value is the robust satisficing policy, and its corresponding uncertainty level is its robustness for the DLSP.



**Fig 3. The solution schemes for the min-max robust decision-making method and the robust satisficing decision-making method.** (A) Solution scheme for the min-max robust decision-making method. (B) Solution scheme for the robust satisficing decision-making method.

doi:10.1371/journal.pone.0166448.g003

The solution scheme is designed based on several lemmas and theorems which will be provided in next section. The solution procedure for Problem 1 can be outlined as follows:

1. The monotonic relationship between the uncertainty level and the worst-case LSP will be established (Lemmas 1 and 2), which supports the division scheme of the uncertainty level. As the uncertainty level increases, the worst-case LSP will decrease such that a critical value of satisfying the DLSP will be reached.

2. The robust satisficing optimality theorems are provided and proved (Theorems 1 and 2), which will help improve the policy towards a higher robustness. A trade-off relationship between the robustness and the DLSP is also proved (Theorem 3).
3. The highest worst-case LSP is calculated as well as the corresponding optimal robust policy based on robust dynamic programming, which will be compared with the DLSP to determine whether the uncertainty level increases or not. And the robust satisficing policy generation algorithm is presented (Algorithm 1).
4. Furthermore, a robustness evaluation algorithm is presented, which can be used to evaluate the robustness of a fixed policy for a given DLSP (Algorithm 2).

### 4 Solution Method

In this section, we will explain each step of the solution procedure for Problem 1 in detail.

First, we begin with a simple claim. It is obvious that the choice of the DLSP will influence the robustness. Thus, a proper DLSP is determined such that a robust satisficing policy with nonzero robustness can be generated.

Let  $V_{\pi}^*(s_p, \alpha) = \min_{P \in U_P(\alpha, \tilde{P})} Pr^{\pi}$  denote the worst-case LSP of a given policy  $\pi$  starting from  $s_p$  at the uncertainty level  $\alpha$ .

**Claim 1:** If the given DLSP  $p_c$  is higher than the LSP of the optimal policy without considering uncertainty, the robustness of all the policies for this DLSP will be zero.

**Proof:** The proof will be provided by contradiction. First, we assume that  $\pi_1$  is the optimal policy without considering uncertainty, and the worst case LSP of  $\pi_1$  is  $V_{\pi_1}^*(s_p, 0) < p_c$ . Assuming that there is an arbitrary policy  $\pi_2$  with robustness  $\alpha_2 \neq 0$  for the DLSP  $p_c$ , we have

$$V_{\pi_2}^*(s_p, \alpha_2) = \min_{P \in U(\alpha_2, \tilde{P})} Pr^{\pi_2} \geq p_c \tag{10}$$

According to the assumption of the optimal policy  $\pi_1$ , the following inequalities can be established

$$V_{\pi_2}^*(s_p, \alpha_2) = \min_{P \in U(\alpha_2, \tilde{P})} Pr^{\pi_2} \leq \min_{P \in U(0, \tilde{P})} Pr^{\pi_2} \leq \min_{P \in U(0, \tilde{P})} Pr^{\pi_1} = V_{\pi_1}^*(s_p, 0) < p_c \tag{11}$$

It's obvious that the relationship of Eq (11) contradicts the assumption in Eq (10). Thus, the arbitrary policy  $\pi_2$  with non-zero robustness  $\pi_2$  does not exist.

According to Claim 1, if the given DLSP is higher than the LSP of the optimal policy without uncertainty, there would be a robust satisficing policy. In this case, the given DLSP cannot be satisfied by any policy, which means the mission specifications cannot be met. Therefore, if we would like to generate a robust satisficing policy, the desired satisfaction probability must be proper. Besides, as defined in the info-gap uncertainty model (4), the maximum robustness lies in [0, 1], which corresponds to a robust satisficing policy with the proper DLSP satisfied. In the subsequent section, it is assumed that the given DLSP is proper, i.e., not higher than the LSP of the optimal policy without considering uncertainty.

**Remark 2:** For a given proper DLSP, the maximum robustness Eq (5) will be achieved by the robust satisficing policy, i.e.,  $\hat{\alpha}(\pi^*(p_c), p_c)$ . The robust satisficing policy  $\pi^*(p_c)$  Eq (7) maximizes  $\hat{\alpha}(\pi, p_c)$ , which is conditional on some  $p_c$ , and makes the condition  $Pr^{\pi^*(p_c)}(s_0 \models \phi) \geq p_c$  guaranteed for any  $P \in U(\hat{\alpha}, \tilde{P})$ .

### 4.1 Monotonicity

The monotonic relationship between the uncertainty level and the worst-case LSP of a given policy is presented.

**Lemma 1 (Monotonicity 1):** For a given policy  $\pi$ , the worst-case LSP does not increase along with the uncertainty level, i.e., if  $\alpha_1 < \alpha_2$ , then  $V_\pi^*(s_p, \alpha_1) \geq V_\pi^*(s_p, \alpha_2)$ .

**Proof:** If the uncertainty level increases from  $\alpha_1$  to  $\alpha_2$ , by the nesting axiom, the uncertainty set  $U_p(\alpha_2, \tilde{P})$  will contain  $U_p(\alpha_1, \tilde{P})$ , i.e.,  $U_p(\alpha_1, \tilde{P}) \subset U_p(\alpha_2, \tilde{P})$ . It is assumed that  $U_p(\alpha_2, \tilde{P}) = U_p(\alpha_1, \tilde{P}) \cup \Omega$ , and  $\Omega$  is a nonempty set. So there will be a relationship of the worst-case LSP between  $\alpha_1$  and  $\alpha_2$  as follows:

$$\begin{aligned} V_\pi^*(s_p, \alpha_2) &= \min_{P_p \in U_p(\alpha_2, \tilde{P})} Pr^\pi = \min_{P_p \in U_p(\alpha_1, \tilde{P}) \cup \Omega} Pr^\pi \\ &= \min\left\{ \min_{P_p \in U_p(\alpha_1, \tilde{P})} Pr^\pi, \min_{P_p \in \Omega} Pr^\pi \right\} \\ &\leq \min_{P_p \in U_p(\alpha_1, \tilde{P})} Pr^\pi = V_\pi^*(s_p, \alpha_1) \end{aligned}$$

It can be concluded that if  $\alpha_1 < \alpha_2$ ,  $V_\pi^*(s_p, \alpha_1) \geq V_\pi^*(s_p, \alpha_2)$ .

According to Lemma 1, the worst-case LSP  $V_\pi^*(s_{p0}, \alpha)$  of a fixed policy monotonically decreases at the uncertainty level  $\alpha$ . So as  $\alpha$  increases, the worst-case LSP of a fixed policy decreases. If  $\alpha$  increases to  $\alpha_m$  such that  $V_\pi^*(s_{p0}, \alpha_m) \geq p_c$ , and  $V_\pi^*(s_{p0}, \alpha_m + \epsilon) < p_c$  ( $\epsilon$  is an infinitely small value), then  $\alpha_m$  is the robustness of the fixed policy that is sought for.

Second, the monotonic relationship between the uncertainty level and the highest worst-case LSP is established.

Let  $H^*(s_p, \alpha) = \max_{\pi \in \Pi} \min_{P_p \in U_p(\alpha, \tilde{P})} Pr^\pi$  denote the highest worst-case LSP at the uncertainty level  $\alpha$  from the initial state  $s_p$ , i.e.,  $H^*(s_p, \alpha) = \max_{\pi \in \Pi} V_\pi^*(s_p, \alpha)$ . The policy with the highest worst-case LSP at a given uncertainty level is called the optimal robust policy at this uncertainty level, which is also referred to as the min-max robust policy in this paper.

**Lemma 2 (Monotonicity 2):** For all the policies, the highest worst-case LSP decreases as the uncertainty level increases, i.e., if  $\alpha_1 < \alpha_2$ , then  $H^*(s_p, \alpha_1) > H^*(s_p, \alpha_2)$ .

**Proof:** It is assumed that  $\pi_1$  is the optimal robust policy at the uncertainty level  $\alpha_1$ ,  $\pi_2$  is the optimal robust policy at the uncertainty level  $\alpha_2$ , and  $\alpha_1 < \alpha_2$ . Since  $\pi_1$  is the optimal robust policy at the uncertainty level  $\alpha_1$ , we have  $H^*(s_p, \alpha_1) = V_{\pi_1}^*(s_p, \alpha_1) > V_{\pi_2}^*(s_p, \alpha_1)$ . Because  $\alpha_1 < \alpha_2$ , we have  $V_{\pi_2}^*(s_p, \alpha_1) > V_{\pi_2}^*(s_p, \alpha_2) = H^*(s_p, \alpha_2)$  according to Lemma 1. It can be concluded that if  $\alpha_1 < \alpha_2$ ,  $H^*(s_p, \alpha_1) > H^*(s_p, \alpha_2)$ .

According to the monotonicity in Lemma 2, there will exist a maximal value of  $\alpha_m$  such that  $H^*(s_{p0}, \alpha_m) \geq p_c$  and  $H^*(s_{p0}, \alpha_m + \epsilon) < p_c$  where  $\epsilon$  is an infinitely small value. This maximal value  $\alpha_m$  is the robustness of the optimal robust policy for the DLSP, and the optimal robust policy corresponding to the uncertainty level  $\alpha_m$  is the robust satisficing policy.

### 4.2 Robust Satisficing Optimality

In order to obtain a robust satisficing policy, the feasible policy should be improved towards the direction with higher robustness. Thus Theorems 1 and 2 are given and proved to define the direction with higher robustness. In Theorem 3, the trade-off relationship between the DLSP and the robustness will be proved.

**Theorem 1 (Robustness optimality):** For a given DLSP  $p_c$ , the policy which maximizes the worst-case LSP will lead to a robustness higher than or equal to that obtained by any other policy at the same uncertainty level.

**Proof:** It is assumed that the robustness of  $\pi_1$  for the desired level  $p_c$  is  $\alpha_1$ , and  $\pi_1$  cannot maximize the worst-case LSP. Thus policy  $\pi_1$  satisfies  $V_{\pi_1}^*(s_{p_0}, \alpha_1) \geq p_c$ , and  $V_{\pi_1}^*(s_{p_0}, \alpha_1 + \epsilon) < p_c$ , where  $\epsilon$  is an infinitely small value.

Let  $\pi_2$  be the policy that can maximize the worst-case LSP at the uncertainty level  $\alpha_1$ . We have

$$V_{\pi_2}^*(s_{p_0}, \alpha_1) = \max_{\pi \in \Pi} \min_{p_p \in U_p(\alpha_1, \hat{p})} Pr^\pi > \min_{p_p \in U_p(\alpha_1, \hat{p})} Pr^{\pi_1} = V_{\pi_1}^*(s_{p_0}, \alpha_1) \geq p_c$$

First the critical case for  $\pi_2$  is considered. It is assumed that for the infinitely small value  $\epsilon$ ,  $V_{\pi_2}^*(s_{p_0}, \alpha_1 + \epsilon) < p_c$ . Then the robustness of  $\pi_2$  is also  $\alpha_1$ . However, in other cases, if  $V_{\pi_2}^*(s_{p_0}, \alpha_1 + \epsilon) \geq p_c$ , there would be  $\alpha_2 \geq \alpha_1 + \epsilon$ , such that  $V_{\pi_2}^*(s_{p_0}, \alpha_2) \geq p_c$  and  $V_{\pi_2}^*(s_{p_0}, \alpha_2 + \epsilon) < p_c$ . In this case, the robustness of  $\pi_2$  is  $\alpha_2$ . Therefore, the robustness of policy  $\pi_2$  is higher than that of policy  $\pi_1$ .

**Theorem 2 (Robust satisficing optimality):** For a given DLSP  $p_c$ , a min-max robust policy which can maximize the worst-case LSP at a certain uncertainty level (maximum robustness) can be found, so it is the robust satisficing policy.

**Proof:** Let  $\pi^*(p_c)$  be the robust satisficing policy with the maximum robustness  $\alpha_m = \hat{\alpha}(\pi^*(p_c), p_c)$ . We have  $V_{\pi^*(p_c)}^*(s_{p_0}, \alpha_m) \geq p_c$  and  $V_{\pi^*(p_c)}^*(s_{p_0}, \alpha_m + \epsilon) < p_c$ , where  $\epsilon$  is an infinitely small value.

It is assumed that  $\pi_{mmm}$  is the min-max robust policy at the uncertainty level  $\alpha_m$  and  $V_{\pi_{mmm}}^*(s_p, \alpha_m)$  is the optimal worst-case LSP, which satisfies

$$V_{\pi_{mmm}}^*(s_{p_0}, \alpha_m) \geq V_{\pi^*(p_c)}^*(s_{p_0}, \alpha_m) \geq p_c$$

First, it is assumed that the inequality holds. According to the proof of Theorem 1, the robustness of  $\pi_{mmm}$  is higher than that of  $\pi^*(p_c)$  in all cases but the critical case. Thus there will exist  $\alpha_n > \alpha_m$ , such that  $V_{\pi_{mmm}}^*(s_{p_0}, \alpha_n) \geq p_c$  and  $V_{\pi_{mmm}}^*(s_{p_0}, \alpha_n + \epsilon) < p_c$ . However, this violates the definition of the robust satisficing policy about maximal robustness. So the inequality condition is false. Under both the critical case and equality conditions, the robustness of  $\pi_{mmm}$  is equal to that of  $\pi^*(p_c)$ . That is, the min-max robust policy at the uncertainty level  $\alpha_m$  is the robust satisficing policy for the DLSP  $p_c$ .

**Theorem 3 (Trade-off Theorem):** If the DLSP decreases, the robustness of a given policy or a robust satisficing policy will be non-decreasing, i.e., if  $p_c > p'_c$ , then  $\hat{\alpha}(\pi, p_c) \leq \hat{\alpha}(\pi, p'_c)$ .

**Proof:** It is assumed that a feasible robustness set for a given policy  $\pi$  is denoted as

$$\Lambda(\pi, p_c) = \left\{ \alpha : \min_{p_p \in U_p(\alpha, \hat{p})} Pr^\pi \geq p_c \right\}$$

Thus the robustness is  $\hat{\alpha}(\pi, p_c) = \max \Lambda(\pi, p_c)$ .

It is assumed that  $p_c > p'_c$ , and the feasible robustness sets  $\Lambda(\pi, p_c)$  and  $\Lambda(\pi, p'_c)$  are non-empty. Let  $\alpha_1 \in \Lambda(\pi, p_c)$ , which means that  $\min_{p_p \in U_p(\alpha_1, \hat{p})} Pr^\pi \geq p_c \geq p'_c$ . It can be concluded that  $\alpha_1 \in \Lambda(\pi, p'_c)$ . Thus we have  $\Lambda(\pi, p_c) \subseteq \Lambda(\pi, p'_c)$ . According to the definition, it follows that  $\hat{\alpha}(\pi, p_c) = \max \Lambda(\pi, p_c) \leq \max \Lambda(\pi, p'_c) = \hat{\alpha}(\pi, p'_c)$ . The trade-off relationship has thus been proved.

Theorem 3 shows that the robustness is monotonically decreasing at the DLSP. In practice, it may help the decision maker to decide a proper desired performance level by taking the robustness to uncertainty into consideration in order to realize a trade-off between the desired performance level and the robustness.

### 4.3 Calculation of the Highest Worst-Case LSP

According to Theorems 1 and 2, for a given DLSP  $p_c$ , the robust satisficing policy can be regarded as the min-max robust policy at a certain uncertainty level. So in this subsection, we will present the calculation method of the highest worst-case LSP as well as the min-max robust policy at a fixed uncertainty level.

**Lemma 3 (Robust Dynamic Programming [22]):** For the robust control problem, the perfect duality holds:

$$\Phi = \max_{\pi \in \Pi} \min_{P \in U(\bar{z}, \bar{p})} Pr^\pi(s_0 \models \phi) = \min_{P \in U(\bar{z}, \bar{p})} \max_{\pi \in \Pi} Pr^\pi(s_0 \models \phi) \tag{12}$$

The optimal value is given by  $\Phi = v^*(s_0)$ , where  $s_0$  is the initial state. For the product IMDP, the value function is the unique limit value of the convergent vector sequence defined by

$$v_{k+1}(s_p) = \max_{a \in A_p} \min_{P_p \in U_p(\bar{z}, \bar{p})} (P_p)^T v_k, \quad k = 1, 2, \dots \tag{13}$$

The optimal min-max robust control policy is obtained as

$$\pi^*(s_p) = \arg \max_{a \in A_p} \min_{P_p \in U_p(\bar{z}, \bar{p})} (P_p)^T v^* \tag{14}$$

For the proof of the detailed processes and convergence of the min-max robust policy, the readers can refer to Theorem 3 in [22].

In our problem, the highest worst-case LSP  $H^*(s_p, \alpha)$  for a fixed uncertainty level  $\alpha$  can be defined as the unique limit value of the following convergent vector sequence based on Lemma 3 (robust dynamic programming)

$$H_{k+1}(s_p, \alpha) = \max_{a \in A_p} \min_{P_p^a \in [\underline{p}, \bar{p}]} (P_p^a)^T H_k(\bullet, \alpha) \tag{15}$$

where  $H_k(\bullet, \alpha)$  is the vector of  $H_k(s_p, \alpha)$  for all  $s_p \in S_p$  at stage  $k$ , and  $\underline{p}$  and  $\bar{p}$  denote the uncertainty interval for the  $i$ th row of  $P_p^a$ . The initial value for  $H_k(s_p, \alpha)$  is given as the LSP related with AMECs in Subsection 3.3.1. Then, the worst-case LSP can be obtained via value iteration, which is described explicitly in Algorithm 1. Note that during each iteration, an inner minimization for finding the optimal transition probability will be computed based on the dual linear programming, which provides the worst-case condition. Refer to [S1 Appendix](#) for detailed procedures.

With the highest worst-case LSP  $H^*(s_p, \alpha_m)$ , the corresponding min-max robust policy  $\pi^*(s_p)$  can be obtained by setting

$$\pi^*(s_p) = \arg \max_{a \in A_p} \min_{P_p^a \in [\underline{p}, \bar{p}]} (P_p^a)^T H^*(\bullet, \alpha_m) \tag{16}$$

### 4.4 Robust Satisficing Policy Generation Algorithm

In this subsection, a robust satisficing policy generation algorithm based on robust dynamic programming is proposed, as shown in Algorithm 1. The uncertainty level  $\alpha$  is divided between 0 and 1 by  $N$  uniform divisions. For a specified value of  $\alpha$ , the range of transition probability can be determined, and the highest worst-case LSP  $H^*(s_p, \alpha)$  can be calculated via value iteration, as well as the corresponding min-max robust policy  $\pi^*$  (Lines 8–19). Then, the highest worst-case LSP is compared with the given DLSP, and if the satisfaction condition holds, we will increase the uncertainty level by  $1/N$  and repeat the above steps until the highest

worst-case LSP reaches the given DLSP (Lines 20–24). An approximation value  $\alpha^*$  of robustness with the accuracy of  $1/N$  can be determined by  $H^*(s_{p_0}, \alpha^*) \geq p_c$  and  $H^*(s_{p_0}, \alpha^* + 1/N) < p_c$ .

**Algorithm 1** Robust Satisficing Policy Generation

**Required:** product IMDP  $P = \langle S_p, A_p, U_p(\alpha, \tilde{P}), R_p, s_{p_0}, L_p, Acc_p \rangle$   
**Required:** the DLSP  $p_c$   
**Ensure:** Robust satisficing policy  $\pi^*$   
**Ensure:** Maximal Robustness  $\alpha^*$

```

1: Step 0: Initialization
2: Step 0.1: Generate AMECs
3: Generate  $(\tilde{S}_p, \tilde{A}_p), B_0, B_P$ 
4: Step 0.2: Initialize LSP
5: for  $s_p \in \tilde{S}_p$  do
6:    $H(s_p, \alpha) = 1$ 
7: end for
8: for  $s_p \in B_0$  do
9:    $H(s_p, \alpha) = 0$ 
10: end for
11: Step 1: Generate the robust satisficing policy and the maximal robustness
12: Step 1.1: Generate the min-max based robust policy
13:  $\alpha \leftarrow 0$ 
14:  $\Delta \leftarrow \infty$ 
15: while  $\Delta \geq \varepsilon$ 
16:   for  $s_p \in S_p \setminus \tilde{S}_p \cup B_0$  do
17:      $Max_P \leftarrow H(\bullet, \alpha)$ 
18:      $H(s_p, \alpha) = \max_{a \in A_p} \min_{p_p^a \in [p, \bar{p}]} (P_p^a)^T H(\bullet, \alpha)$ 
19:      $\pi = \operatorname{argmax}_{a \in A_p} \min_{p_p^a \in [p, \bar{p}]} (P_p^a)^T H(\bullet, \alpha)$ 
20:      $\Delta = \min(\|H(\bullet, \alpha) - Max_P\|, \Delta)$ 
21:     if  $\Delta \leq \varepsilon$  then
22:        $H^*(s_p, \alpha) = H(s_p, \alpha)$ 
23:     end if
24:   end for
25:   Step 1.2: Update the robustness
26:   if  $\|H^*(s_p, \alpha) - p_c\| \geq 0$  do
27:      $\pi^* \leftarrow \pi$ 
28:      $\alpha^* \leftarrow \alpha + 1/N$ 
29:      $\alpha \leftarrow \alpha + 1/N$ 
30:   goto line 15
31: end while
32: return  $\alpha^*, \pi^*$ 

```

In the following, the convergence of the above algorithm will be provided.

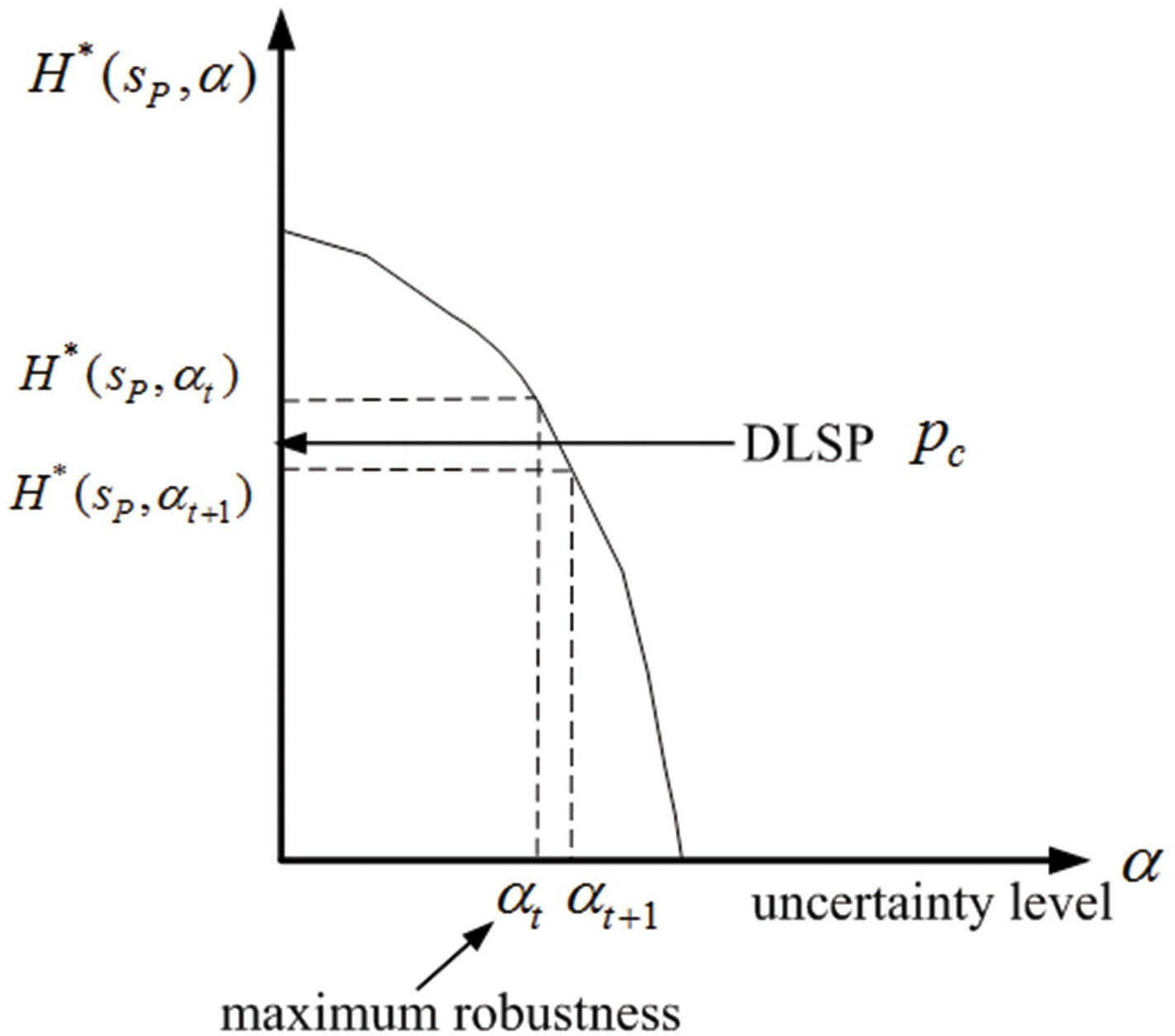
**Theorem 4 (Convergence):** For Problem 1, a robust satisficing policy, which is generated by Algorithm 1, can converge to the maximal robustness with a given proper DLSP  $p_c$  satisfied.

**Proof:** First, according to Lemma 3, for a specified uncertainty level  $\alpha$ , a min-max robust policy can converge to a unique optimal value function, i.e., the highest worst-case LSP. In Algorithm 1, we use the discretized method to divide the uncertainty level into  $N$  equally spaced values  $\alpha_t, t = 1, \dots, N$  between 0 and 1. For each  $\alpha_t$ , the robust policy, generated by the robust dynamic programming, can converge to the highest worst-case LSP  $H^*(s_p, \alpha_t)$ .

Then, due to the monotonicity in Lemma 2, the highest worst-case LSP  $H^*(s_p, \alpha)$  decreases as the uncertainty level  $\alpha$  increases, as shown in Fig 4. Thus we can find a  $\alpha_t$ , such that  $H^*(s_p, \alpha_t) \geq p_c$  and  $H^*(s_p, \alpha_{t+1}) < p_c$ , where  $H^*(s_p, \alpha_t)$  can be achieved by the min-max policy  $\pi_{mm}^t$  at



### highest worst-case LSP



**Fig 4. The monotonicity relationship between the highest worst-case LSP and the uncertainty level, and the maximum robustness.**

doi:10.1371/journal.pone.0166448.g004

the uncertainty level  $\alpha_t$ . In this case,  $\alpha_t$  is the maximum robustness with the accuracy  $1/N$ , as shown in Fig 4.

Finally, according to Theorem 2, the min-max robust policy  $\pi_{mm}^t$  with the highest worst-case LSP  $H^*(s_P, \alpha_t)$  at  $\alpha_t$  is the robust satisficing policy for the DLSP  $p_c$  with the maximum robustness  $\alpha_t$ .

**Remark 3:** The complexity of Algorithm 1 is affected by the size of the product IMDP. The size of the DRA  $|Q|$  is in the worst case, doubly exponential with respect to the LTL formula [39]. The size of the product IMDP is at least  $n = |S| \times |Q|$ , without considering the uncertain transition matrices. For the uncertain IMDP, the complexity of solving the worst-case problem is  $O(n \log(n))$ . And the maximal end components can be generated in  $O(n^2)$  at most. Algorithm 1 spends at most  $O(n^2 a N (\log(1/\epsilon))^2)$  to obtain the robust satisficing policy, where  $N$  is the number of robustness divisions and  $a = |A|$  is the number of actions.

### 4.5 Robustness Evaluation Algorithm

Algorithm 1 is used to generate a robust satisficing policy. However, if there are several feasible policies, a robustness evaluation algorithm is required to choose the policy with the highest robustness. The robustness evaluation problem, for evaluating the robustness of a fixed policy  $\pi$  for a given DLSP  $p_c$ , can be formulated as

$$\hat{\alpha}(\pi, p_c) = \max \left\{ \alpha : \min_{P_p \in U_p(\alpha, \bar{P})} Pr^\pi(s_0 \models \phi) \geq p_c \right\} = \max \{ \alpha : V_\pi^*(s_{p_0}, \alpha) \geq p_c \} \quad (17)$$

The solution scheme of robustness evaluation is similar to that for generating the robust satisficing policy. The Bellman recursion will be used to calculate the worst-case LSP, which satisfies the following conditions

$$V_\pi^*(s_p, \alpha) = \min_{P_p^a \in [P, \bar{P}]} P_p^a(s_p, s'_p)^T V_\pi^*(s'_p, \alpha), s_p \in S_p \quad (18)$$

where  $a = \pi(s_p)$ .  $V_\pi^*(s_p, \alpha)$  for policy  $\pi$  is the unique fixed point of the contraction mapping  $\Phi(\pi, s_p) : \mathbb{R}^{|S_p|} \rightarrow \mathbb{R}^{|S_p|}$  defined by

$$\Phi(\pi, s_p) = \min_{P_p^a \in [P, \bar{P}]} (P_p^a)^T V_\pi(\bullet, \alpha) \quad (19)$$

where  $V_\pi(\bullet, \alpha)$  is the vector of  $V_\pi^*(s_p, \alpha)$  for all  $s_p \in S_p$ . For the proof of the contraction mapping, refer to Lemma 2 in [27]. The worst-case LSP, as the fixed point of the contraction mapping, can be obtained via value iteration. Then a robustness evaluation algorithm is proposed so as to find the robustness of a given policy for a given DLSP, as shown in Algorithm 2.

**Algorithm2** Robustness Evaluation for a Specified Policy

**Required:** product IMDP  $P = \langle S_p, A_p, U_p(\alpha, \bar{P}), R_p, s_{p_0}, L_p, Acc_p \rangle$

**Required:** a specified policy  $\pi$

**Required:** the DLSP  $p_c$

**Ensure:** Robustness  $\alpha^*$

▷ Step 0: Initialization

- 1: Generate AMECs  $(\bar{S}_p, \bar{A}_p), B_0, B_p$
- 2: for  $s_p \in \bar{S}_p \cup B_0$ , initialize  $V_\pi(\bullet, \alpha)$
- 3:  $\Delta \leftarrow \infty$
- 4:  $\alpha \leftarrow 0$
- ▷ Step 1: Evaluate the robustness
- 5: **while**  $V_\pi(s_{p_0}, \alpha) - p_c \geq 0$
- ▷ Step 1.1: Compute the worst-case LSP
- 6: **while**  $\Delta \geq \epsilon$
- 7: **for**  $s_p \in S_p \setminus \bar{S}_p \cup B_0$  **do**
- 8:  $Min_P \leftarrow V_\pi(\bullet, \alpha)$
- 9:  $V_\pi(s_p, \alpha) = \min_{P_p^{\pi(s_p)} \in [P, \bar{P}]} (P_p^{\pi(s_p)})^T V_\pi(\bullet, \alpha)$
- 10:  $\Delta = \min(\|V_\pi(\bullet, \alpha) - Min_P\|, \Delta)$
- 11: **if**  $\Delta \leq \epsilon$  **then**

```

12:          $V_{\pi}^*(s_p, \alpha) = V_{\pi}(s_p, \alpha)$ 
13:     end if
14: end for
15: end while
    ▷ Step 1.2: Update the robustness
16:      $\alpha \leftarrow \alpha + 1/N$ 
17:      $\Delta \leftarrow \infty$ 
18: end while
19: return  $\alpha^* \leftarrow \alpha - 1/N$ 

```

Having formulated the problem and designed the solution algorithms, we will illustrate our methods in the next section.

## 5 Empirical Evaluation

### 5.1 Construction

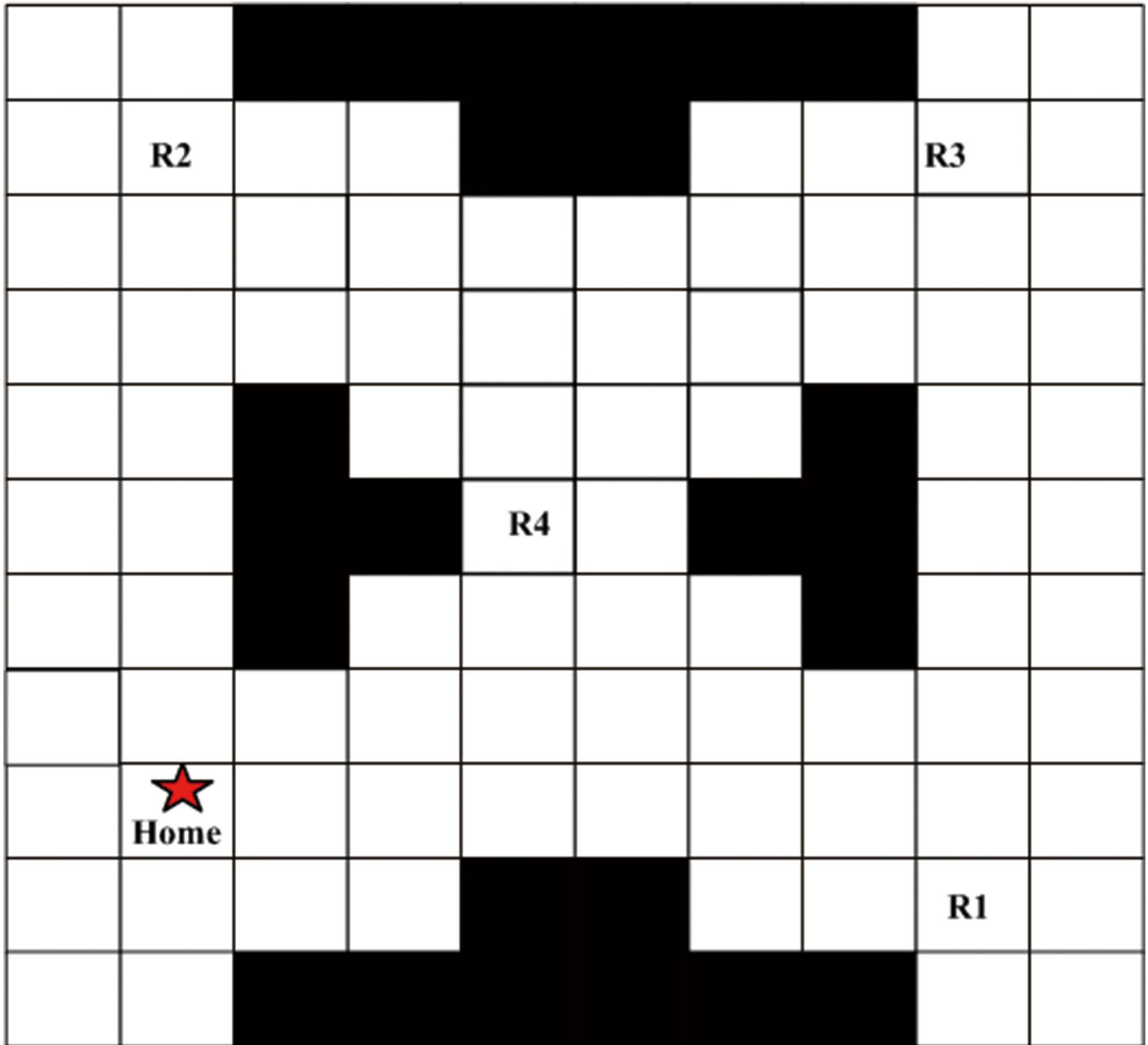
We demonstrate our algorithms on an example of UAV search missions, in which the task for the UAV is to sequentially visit several regions of interest to collect information while always remaining safe. Once the UAV has visited the regions, it should return to the starting point. Simulation experiments are performed in a warehouse, as shown in Fig 5. The LTL formula for this task is  $\phi = G\text{-unsafe} \wedge F((R1 \vee R2) \wedge XF(R3 \wedge XF(R4 \wedge XF\text{home})))$ .

Without any loss of generality, the occupancy grid map is utilized to discretize the workspace, which consists of  $10 \times 11$  cells. Each cell is a square region  $[0, 1] \times [0, 1]$ . Considering a UAV moving in the discrete grid map consisting of  $10 \times 11$  cells, we create a finite MDP abstraction of the UAV system. The UAV is located at one of the cells of the grid map each time, so each cell represents a state of the MDP. The actions enabled at each state are {up, down, left, right}, each of which can direct the UAV to move to one of its three forward adjacent regions. The atomic propositions are {home, R1, R2, R3, R4, unsafe}, each of which labels the states of the MDP with an atomic proposition, as shown in Fig 5. The first four propositions are labeled on the corresponding cells of the grid map, and the unsafe regions are represented by the black blocks in the cells. Due to the symmetry of the region and the UAV, we only need to calculate the estimated transition probability for one region and can apply it to other regions. Through Monte Carlo simulation [27], the transition probabilities for taking action ‘up’ are obtained (*left-forward, forward, right-forward*) = (0.162, 0.687, 0.151), as shown in Fig 6. The transition probabilities for taking other actions are defined in a similar manner. Note that the estimated transition probabilities are not exact as a result of statistical errors. And some undesired state transitions may be caused by control errors or environmental disturbance.

Computations are performed on a 2.6GHz Intel Core i7 processor with 8 GB memory. All the experiments are implemented in the MATLAB. The LTL formula  $\phi$  is transformed into a DRA with 7 states through the ‘ltl2dstar’ software [39], so we can obtain the product IMDP with 770 states by constructing the Cartesian product of the DRA and the IMDP. It takes 0.31s to generate AMECs for the product IMDP.

### 5.2 Implementation and Results

In the following, we will demonstrate our robust satisficing decision-making method in two parts. In the first part, our method is used to generate a robust satisficing policy for the given DLSP, and the effectiveness of this policy is verified by determining whether the DLSP is met. Besides, the trade-off property between robustness and the DLSP is analyzed. In the second and third parts, our method is compared with the min-max robust decision-making method

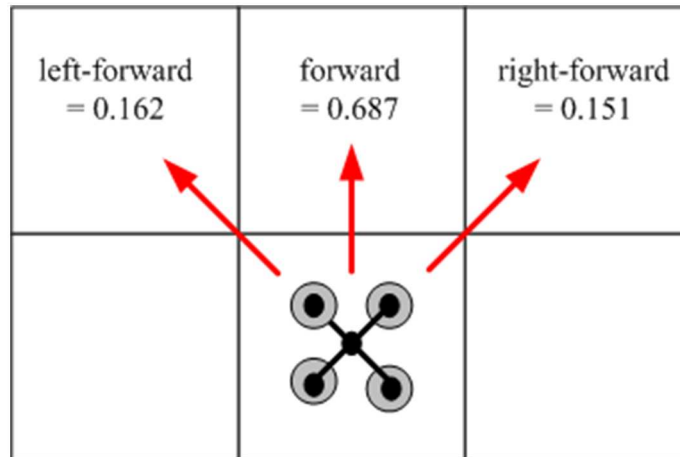


**Fig 5. The mission space with atomic propositions.**

doi:10.1371/journal.pone.0166448.g005

and the robust decision making [37] to illustrate the advantages of the robust satisficing policy in practice.

**5.2.1 Generation and Properties of the Robust Satisficing Policy.** In this subsection, we will verify the effectiveness of the proposed algorithm. To begin with, the simulation parameters are set as follows. For the DLSP, in some applications it might derive from established standards or rules within an organization. In other cases, the DLSP might derive from a combination of the decision maker’s values and understanding of what is possible. In this paper, we take the second approach, and the DLSP  $p_c$  is initially set to be 0.9. And the uncertainty



**Fig 6. The transition probability for action ‘up’.**

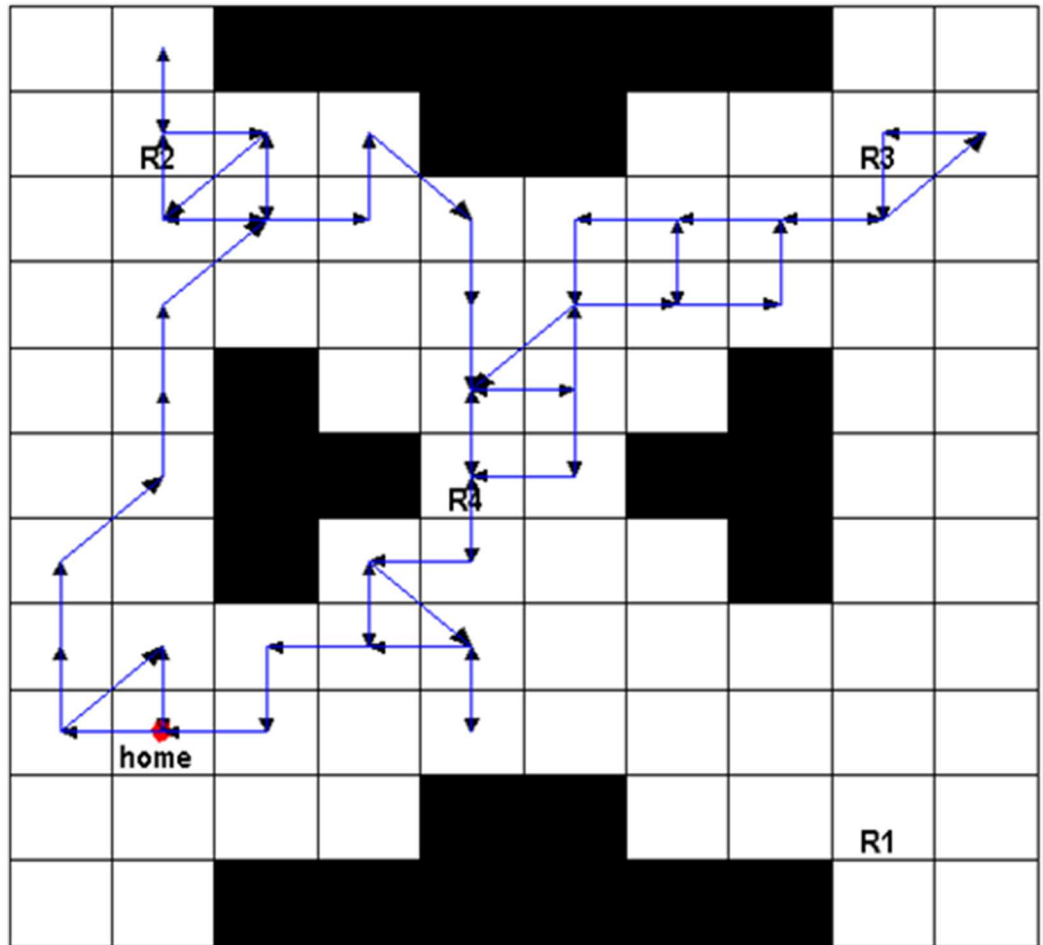
doi:10.1371/journal.pone.0166448.g006

division  $N$  is set to be 100 such that the accuracy of the robustness is 0.01, which is small enough for the transition probability.

Next, a robust satisficing policy is generated by implementing Algorithm 1 in the UAV search mission environment. Aiming at an intuitively expressive form of the resulting policy, we carry out the forward simulation to produce a trajectory of the policy. The UAV starts from the initial state, and determines the action to be taken at the current state according to the generated policy. Then, the transition to the next state is conducted by the worst-case transition probability which can be calculated by the linear programming function (linprog) in the MATLAB. The above procedures are repeated until it reaches a terminal state (AMEC). After the Algorithm 1 is run, the robust satisficing policy  $\pi^*$  is generated in 452.8s with a robustness of 0.40. Through the forward simulation, a trajectory of the UAV as shown in Fig 7 can be generated. The star sign represents the start point, the trajectory is drawn with the blue line, and the arrows represent the direction of motion. There are several overlaps of the path segments as shown in Fig 7, which may cause a little fuzziness, so another figure is drawn to remove the overlaps, as shown in Fig 8. The red path segments represent the influence of uncertainty which makes the motions of UAV deviate from the desired direction of action, while our robust satisficing policy can drive the UAV back.

In order to describe the trade-off relationship between the DLSP and the robustness, we first consider the effect of different DLSPs on the robustness of the generated robust satisficing policy  $\pi^*$ . The robustness of the generated policy  $\pi^*$  for all the DLSPs from 0.0 to 1.0 is calculated by implementing Algorithm 2. And a robustness curve of  $\pi^*$ , representing its varying robustness corresponding to different DLSPs, is shown in Fig 9. It can be seen that the robustness of  $\pi^*$  will decrease as the DLSP increases. Further, we demonstrate the effect of DLSP on the maximum robustness. Considering different DLSPs, we run Algorithm 1 to generate the robust satisficing policy and the corresponding maximum robustness. Fig 10 shows the maximum robustness curve, which represents the relationship between the DLSP and the maximum robustness. And the robustness of  $\pi^*$  and the maximum robustness for different DLSPs are shown in Table 1.

From Fig 10, it can be seen that the slope of the DLSP vs. maximum robustness is always negative. This negative slope represents the trade-off in Theorem 3: as the DLSP increases, the robustness to uncertainty naturally decreases. Note that if the DLSP decreases from 1 to 0.8, the robustness will increase sharply from 0 to around 0.6. Yet, if the DLSP continuously



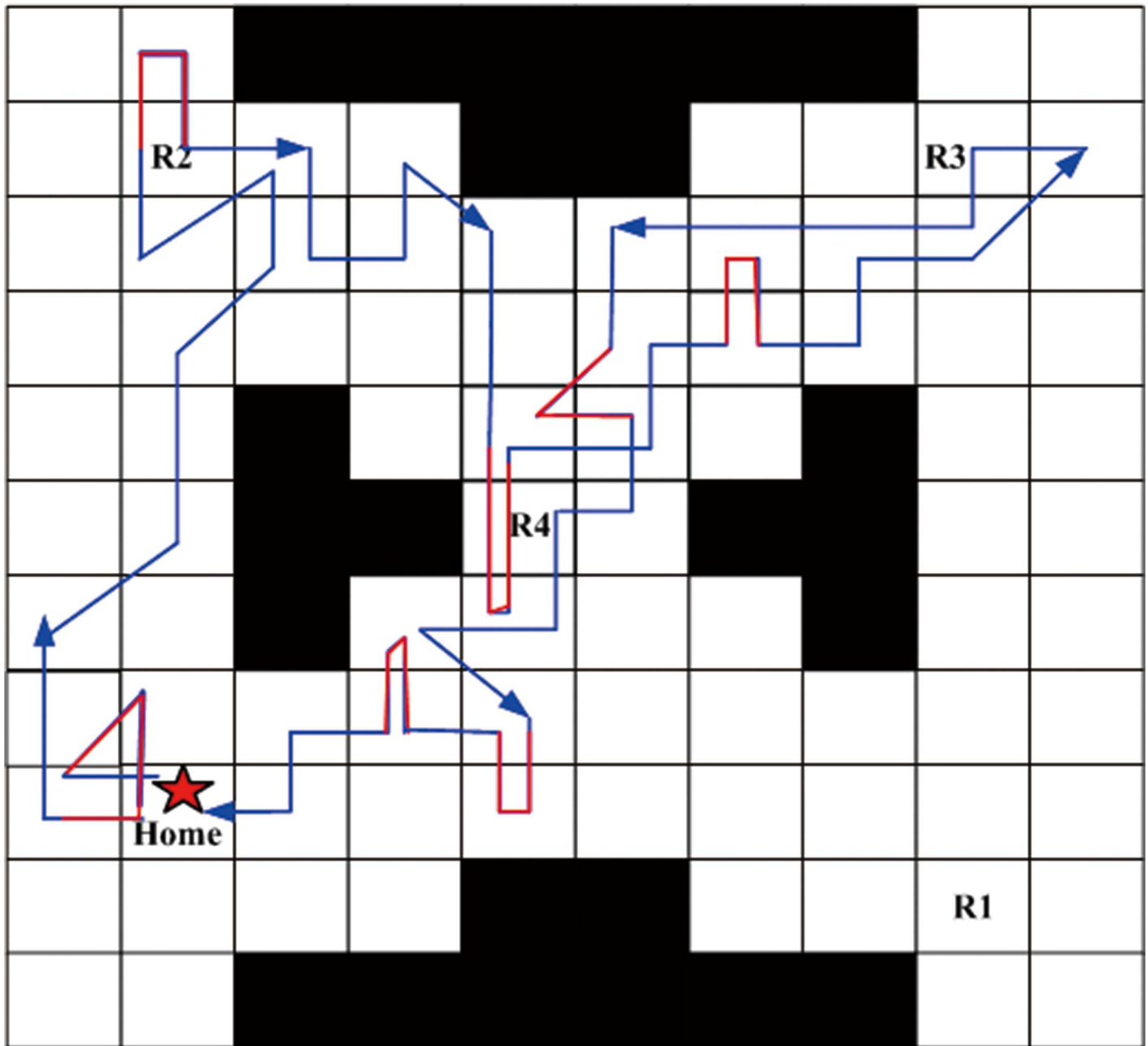
**Fig 7. The original trajectory of the robust satisficing policy through forward simulation.**

doi:10.1371/journal.pone.0166448.g007

decreases, the increase of the robustness will not be remarkable. Thus we recommend a more appropriate DLSP (around 0.85) to the UAV operator instead of 0.9, which can be explained by the principle that a little compromise on performance can lead to a relatively obvious increase in robustness.

Second, we demonstrate that the robust satisficing policy at the maximum uncertainty level (i.e., robustness), will always guarantee the DLSP satisfied. Through Monte Carlo simulation, the true LSP of a resulting policy is computed as the mission success rate over 1000 independent forward simulation runs. Four robust satisficing policies ( $\pi_1, \dots, \pi_4$ ) with robustness ( $\alpha_1, \dots, \alpha_4$ ) for four different DLSPs (0.9, 0.8, 0.7, 0.6) are generated, respectively. For each of these policies, the true LSP is calculated through Monte Carlo simulation and will be compared with its corresponding DLSP. The results are shown in Table 2. The first row shows the DLSP  $p_{\zeta}$  in the second row are the robust satisficing policies, with the robustness in the third row; and in the last row are the true LSPs obtained through Monte Carlo simulations. It can be observed that for each of these policies, the true LSP is higher than its corresponding DLSP. So the robust satisficing policies generated by Algorithm 2 can guarantee the DLSP satisfied.

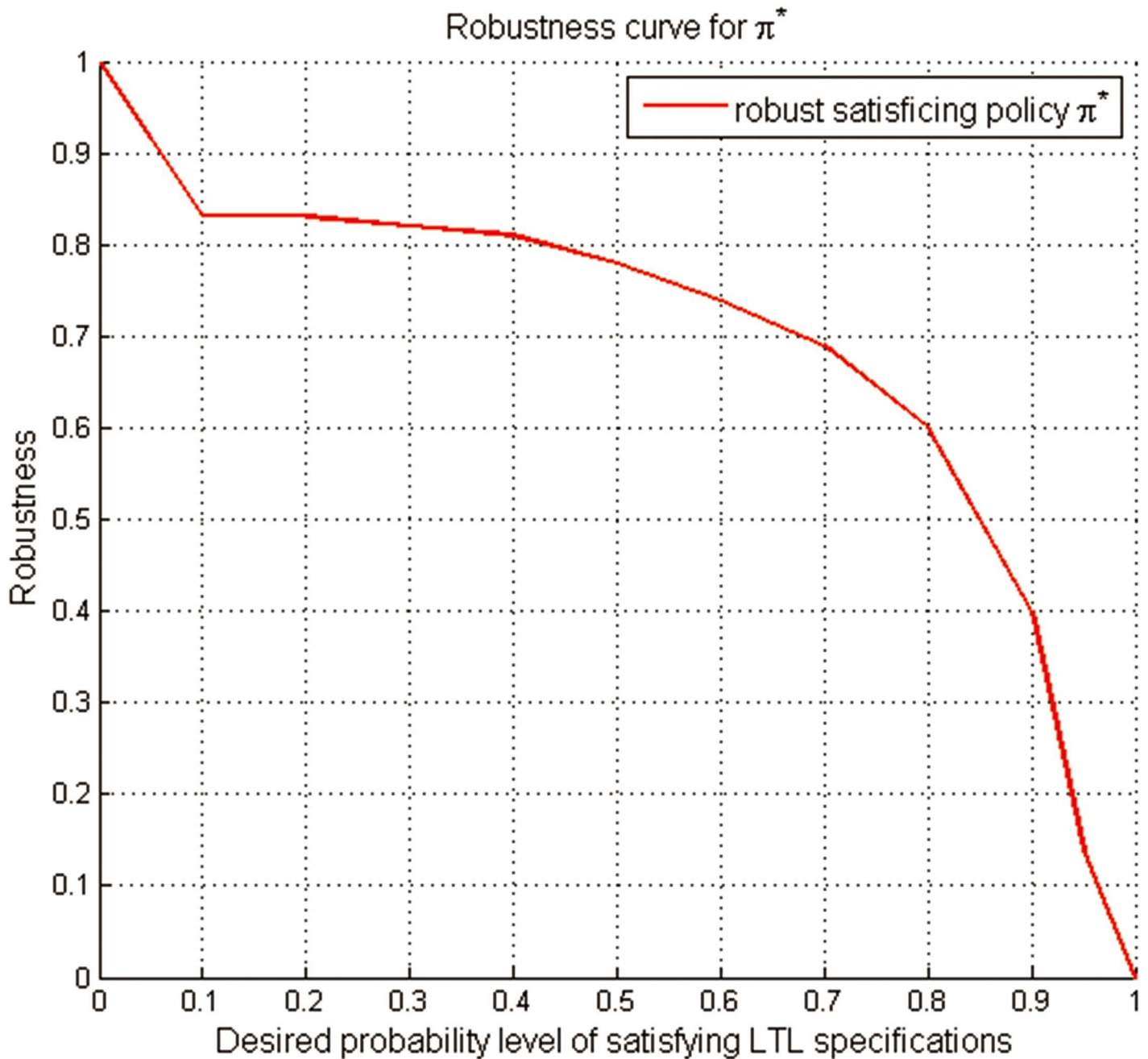
**5.2.2 Comparison with the Min-Max Robust Decision-Making Method.** In this subsection, we will compare our robust satisficing decision-making method with the min-max robust decision-making method [27] in real UAV applications. It is assumed that the desired mission



**Fig 8. The trajectory of the robust satisficing policy removing overlaps.**

doi:10.1371/journal.pone.0166448.g008

success rate (i.e., DLSP) is  $p_c = 0.85$  for the UAV, which is the trade-off performance obtained from the robustness curve. The environment and mission specifications are set identically in Subsection 5.1. The estimated uncertainty level of transition probability is initially assumed to be 0.2. By using the min-max robust decision-making method, a robust policy  $\pi_m$  can be generated at the estimated uncertainty level, and the trajectory generated by forward simulations is shown in Fig 11. Then our robust satisficing decision-making method is used to generate a robust satisficing policy  $\pi_s$  for the DLSP  $p_c = 0.85$ . The resulting policy  $\pi_s$  has a robustness of 0.54, and the trajectory is shown in Fig 12. The true LSPs for these two policies are calculated



**Fig 9. The robustness curve for the robust satisficing policy  $\pi^*$ .**

doi:10.1371/journal.pone.0166448.g009

through Monte Carlo simulations, and the values are 0.97 and 0.96 for  $\pi_m$  and  $\pi_s$ , respectively. It can be seen that if the true uncertainty level is 0.2, both  $\pi_m$  and  $\pi_s$  satisfy the DLSP. However, in real applications, the estimated uncertainty level may be inaccurate as a result of environment disturbance. In order to determine the failure boundaries for  $\pi_m$  and  $\pi_s$ , their true LSPs are evaluated at different uncertainty levels from 0.1 to 1, and the results are shown as the histograms and curves in Figs 13 and 14, as well as the data in Table 3.



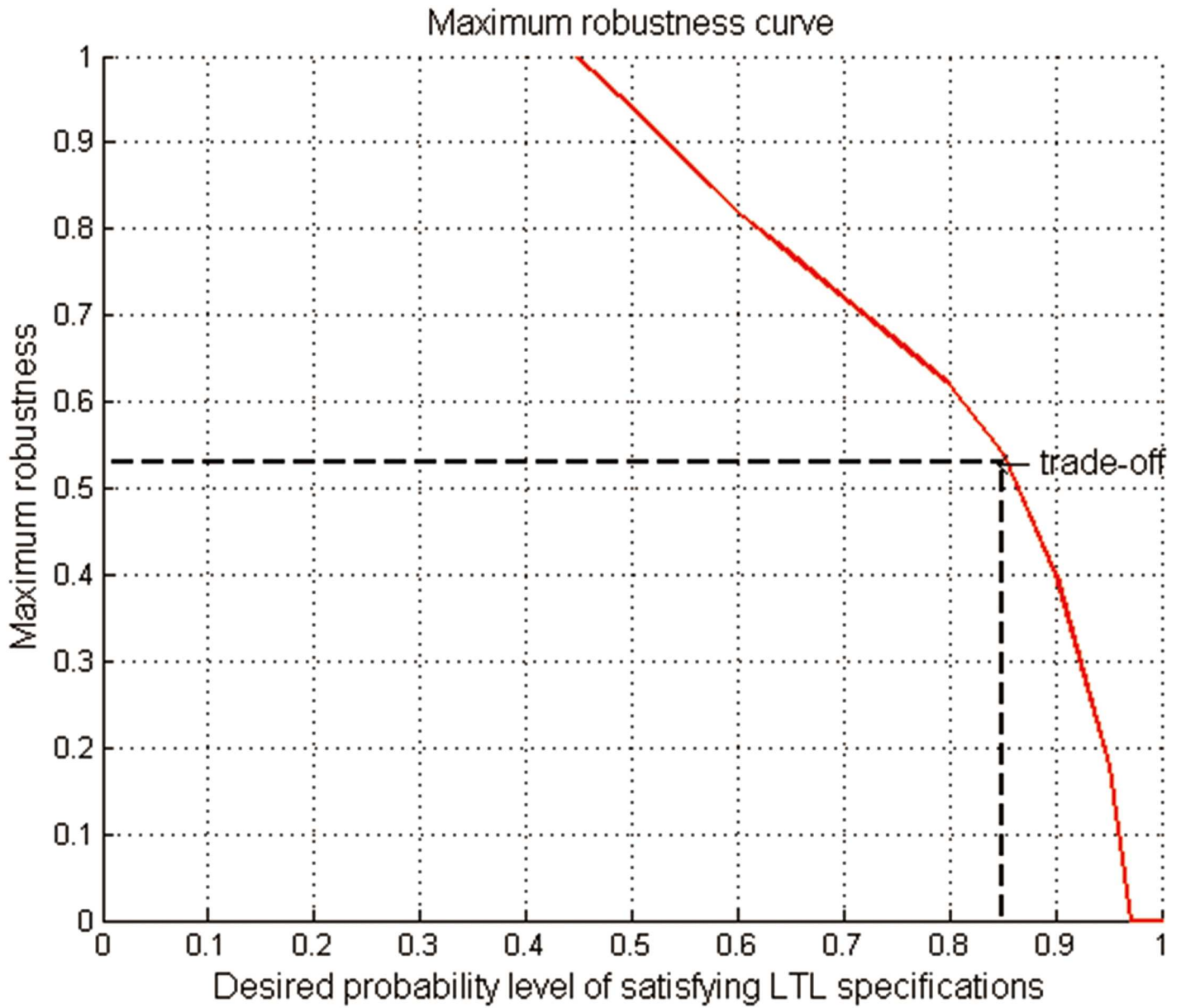


Fig 10. The maximum robustness curve.

doi:10.1371/journal.pone.0166448.g010

Table 1. Robustness for different desired probability levels of satisfying LTL specifications.

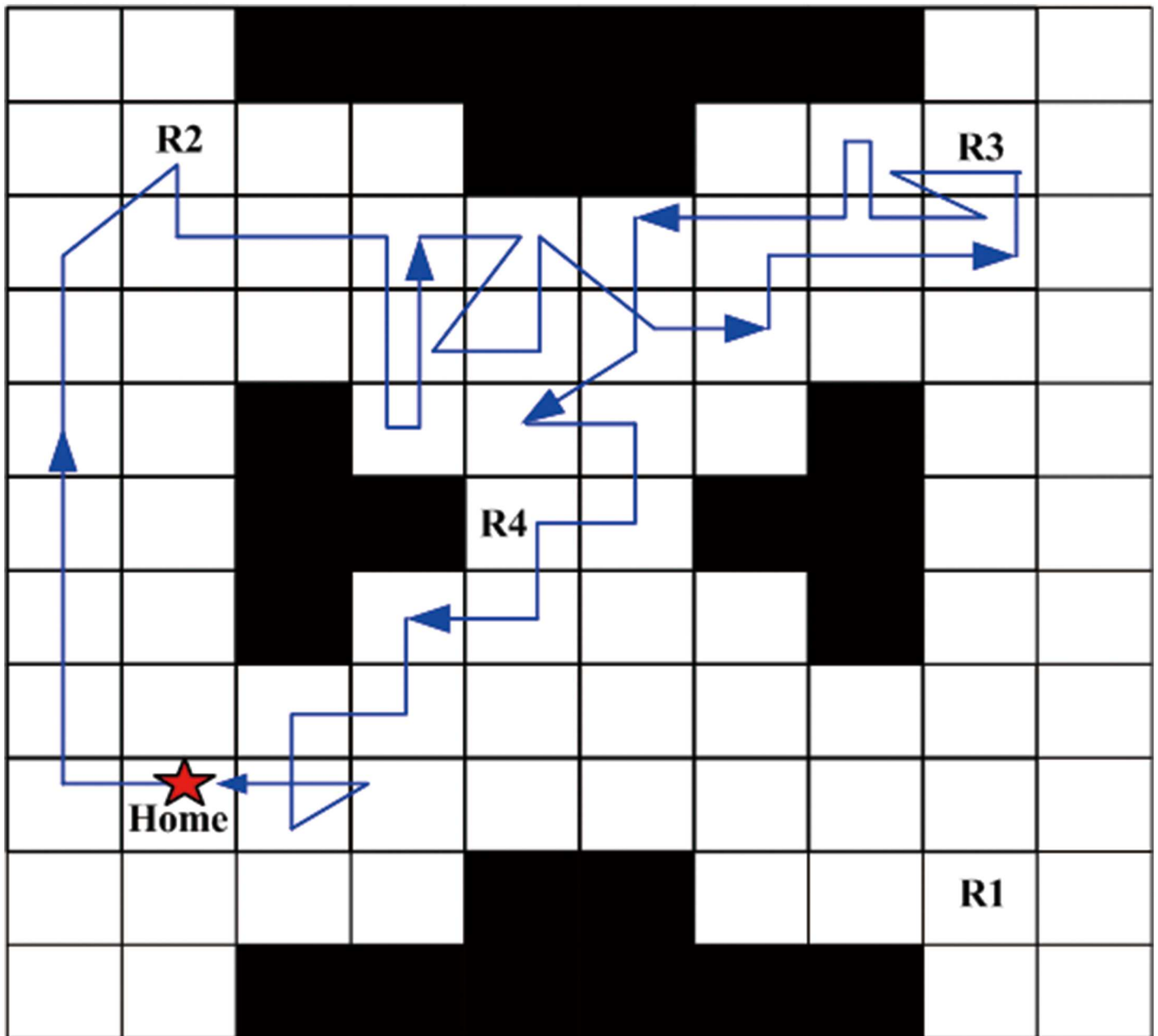
DLSP $p_c$	0.00	0.10	0.20	0.30	0.40	0.50
robustness of $\pi^*$	1.00	0.83	0.83	0.82	0.81	0.78
maximum robustness	1.00	1.00	1.00	1.00	1.00	0.94
DLSP $p_c$	0.60	0.70	0.80	0.90	0.95	1.00
robustness of $\pi^*$	0.74	0.69	0.60	0.40	0.14	0.00
maximum robustness	0.82	0.72	0.62	0.40	0.18	0.00

doi:10.1371/journal.pone.0166448.t001

**Table 2. Desired probability level vs. true probability of satisfying the LTL specifications.**

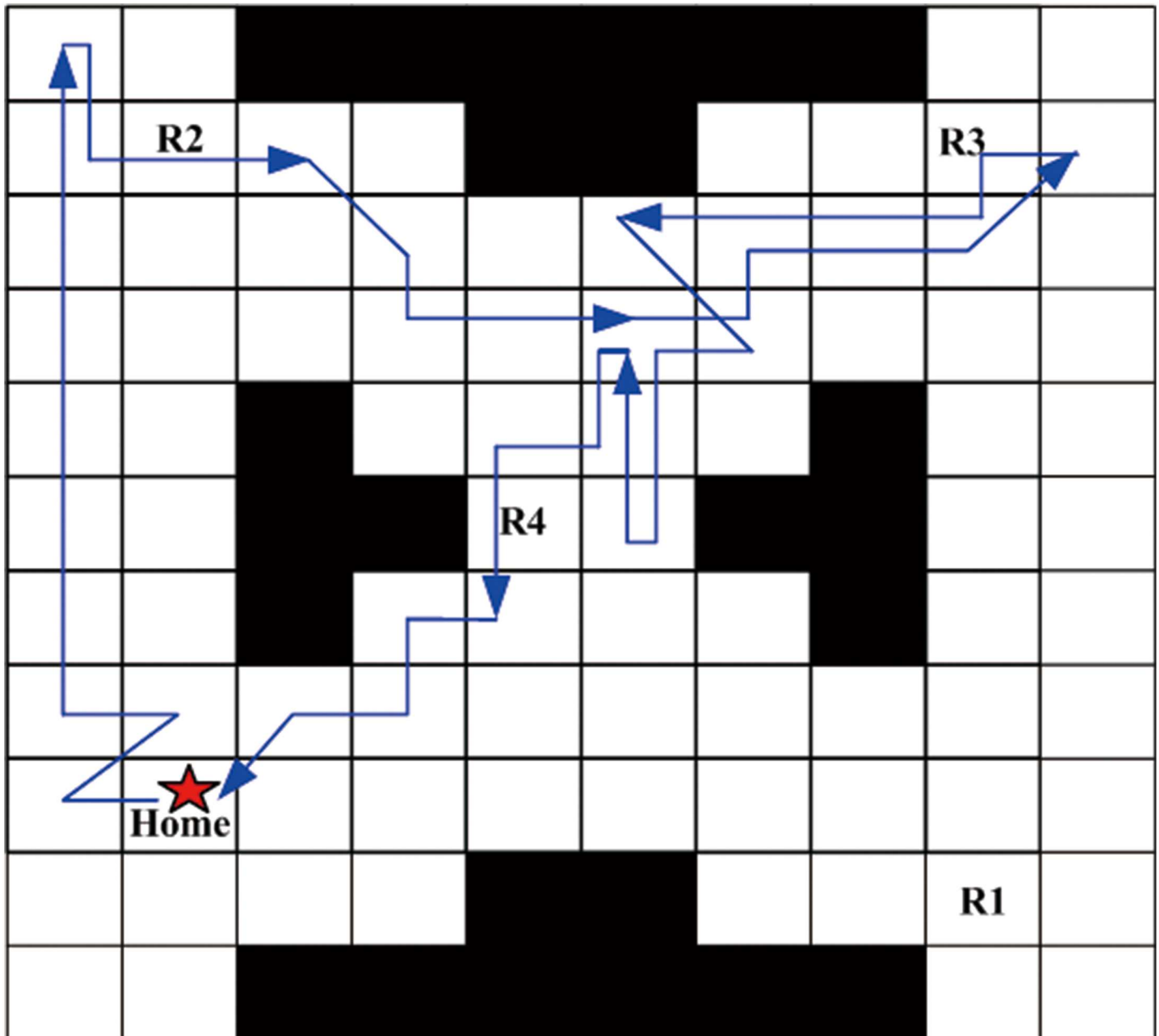
DLSP $p_c$	0.90	0.80	0.70	0.60
robust satisficing policy	$\pi_1$	$\pi_2$	$\pi_3$	$\pi_4$
robustness	0.40	0.62	0.72	0.82
true LSP	0.955	0.947	0.788	0.613

doi:10.1371/journal.pone.0166448.t002



**Fig 11. The trajectory of the min-max robust policy  $\pi_m$  through forward simulations.**

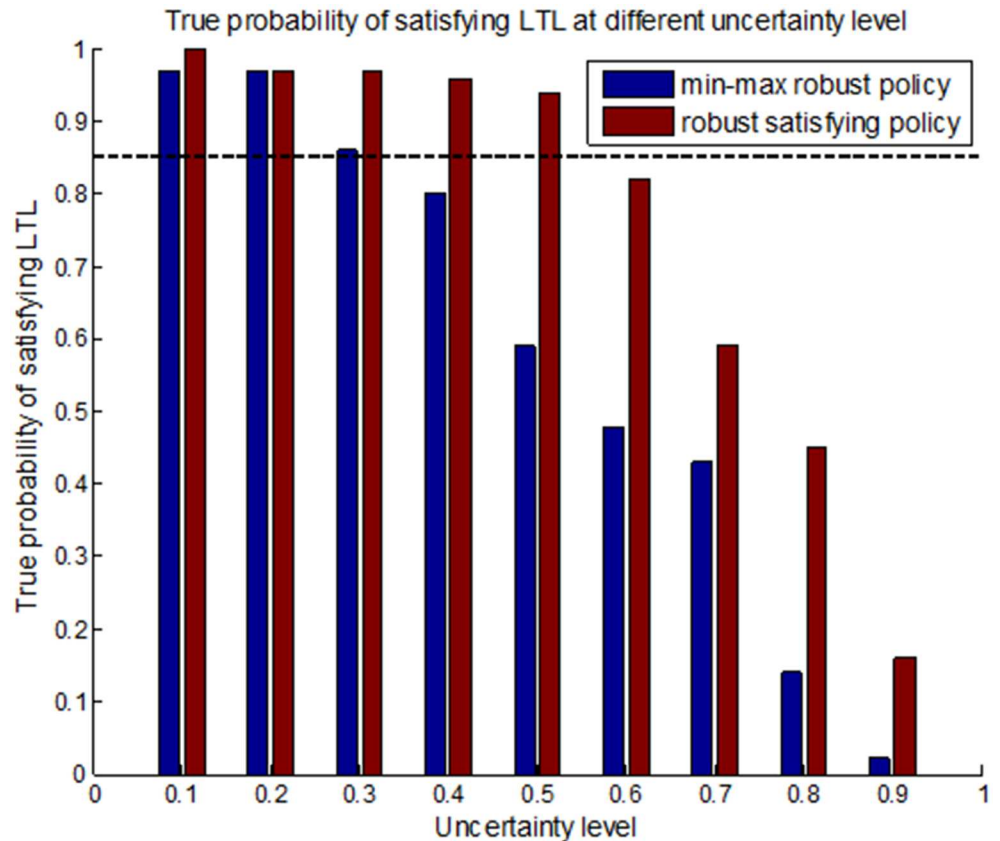
doi:10.1371/journal.pone.0166448.g011



**Fig 12. The trajectory of the robust satisficing policy  $\pi_s$  through forward simulations.**

doi:10.1371/journal.pone.0166448.g012

From Figs 13 and 14 and Table 3, it can be seen that if the uncertainty level exceeds 0.37, the min-max robust policy  $\pi_m$  will lead to a value smaller than 0.85 such that it cannot guarantee the DLSP. However, for the robust satisficing policy  $\pi_s$ , the DLSP will be met until the uncertainty level extends to 0.54. And from the curves, it can be observed that the true LSP of the robust satisficing policy is higher than that of the min-max robust policy at all the uncertainty levels except from 0.15 to 0.21. Therefore, it can be concluded that our robust satisficing policy can tolerate higher uncertainty than the min-max robust policy, and is much more effective in real applications to guarantee a desired performance level.



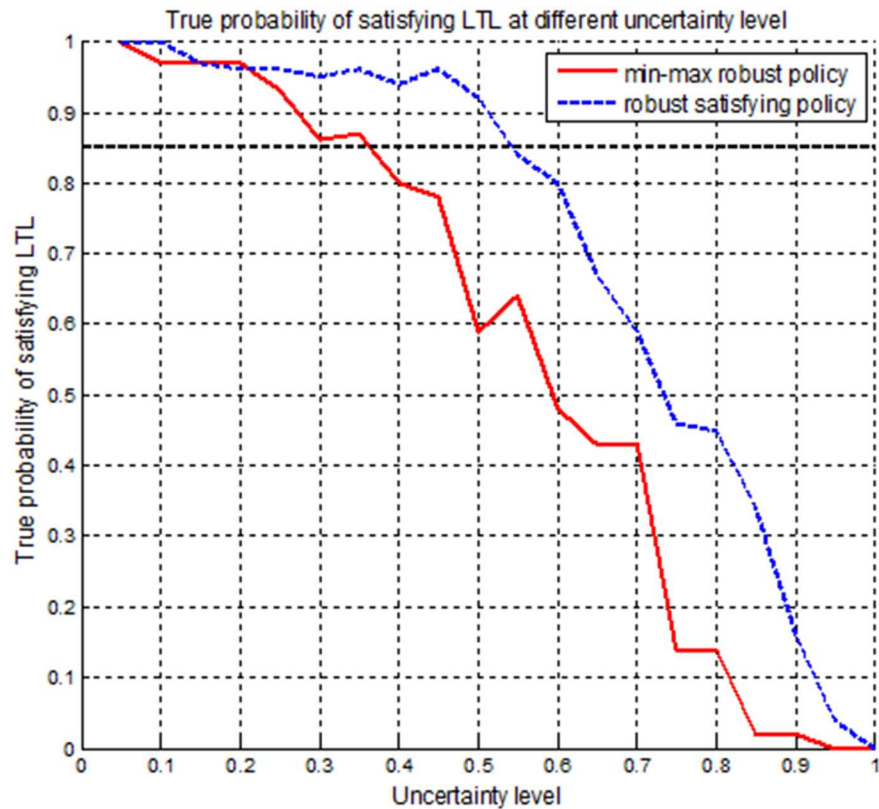
**Fig 13. The histograms representing the true LSPs for the min-max robust policy  $\pi_m$  and the robust satisfying policy  $\pi_s$  at different uncertainty levels.**

doi:10.1371/journal.pone.0166448.g013

**5.2.3 Comparison with Robust Decision Making.** Considering the robustness optimization problem under severe uncertainty, we compare the proposed method with the robust decision making (RDM) [37, 40], which also provides a structured approach to making robust decisions under severe uncertainty. The RDM employs two decision criteria, i.e., definitions of robustness, adopted in the risk analysis literature [40, 41]. The first defines a robust policy as one that trades some optimal performance for less sensitivity to uncertainty, namely the limited degree of confidence (LDC) criterion. And the second defines a robust policy as one that performs relatively well compared with the alternatives over a wide range of futures, namely the wide range of futures (WRF).

With respect to the above two robustness criteria, two RDM robust policies  $\pi_{r1}$  and  $\pi_{r2}$  are generated by the robust decision making [41], and the trajectories are shown in Figs 15 and 16. In order to compare the robust satisfying policy  $\pi_s$  (generated by our method for the DLSP  $p_c = 0.85$  in the previous subsection) with the RDM robust policies  $\pi_{r1}$  and  $\pi_{r2}$ , their true LSPs are evaluated at different uncertainty levels from 0.1 to 1, and the results are shown as the curves in Fig 17 and the data in Table 4.

From Fig 17 and Table 4, it can be seen that the RDM robust policy  $\pi_{r1}$  performs almost the same as the robust satisfying policy  $\pi_s$ , since it defines robustness as trading some optimal performance for less sensitivity to uncertainty, which is similar to the info-gap decision theory. The RDM robust policy  $\pi_{r2}$  has the lowest LSP when the range of uncertainty level is from 0 to



**Fig 14. The curves representing the true LSPs for the min-max robust policy  $\pi_m$  and the robust satisficing policy  $\pi_s$  at different uncertainty levels.**

doi:10.1371/journal.pone.0166448.g014

0.6, and when the uncertainty level increases above 0.7, it will have the highest LSP. The RDM robust policy  $\pi_{r2}$  performs relatively well over all uncertainty levels, for it defines robustness as performing relatively well over a wide range of plausible futures. Our method begins with the estimated value for each uncertain system input and then sequentially chooses values increasingly farther away from the expected inputs, while the RDM samples all the possible uncertain parameters to identify the conditions of system failure. In this problem, it is not difficult to obtain an estimated value of the uncertain transition probability via Monte Carlo simulations. Thus, with the estimated value and a predefined performance level, our method will be more applicable. Given the diversity of definitions of robustness, and the differing judgments called for in implementing alternative robust decision methods, it is perhaps surprising they often reach similar results. Therefore, our method provides an alternative way to make robust decisions to handle the severe uncertainty, like the RDM method.

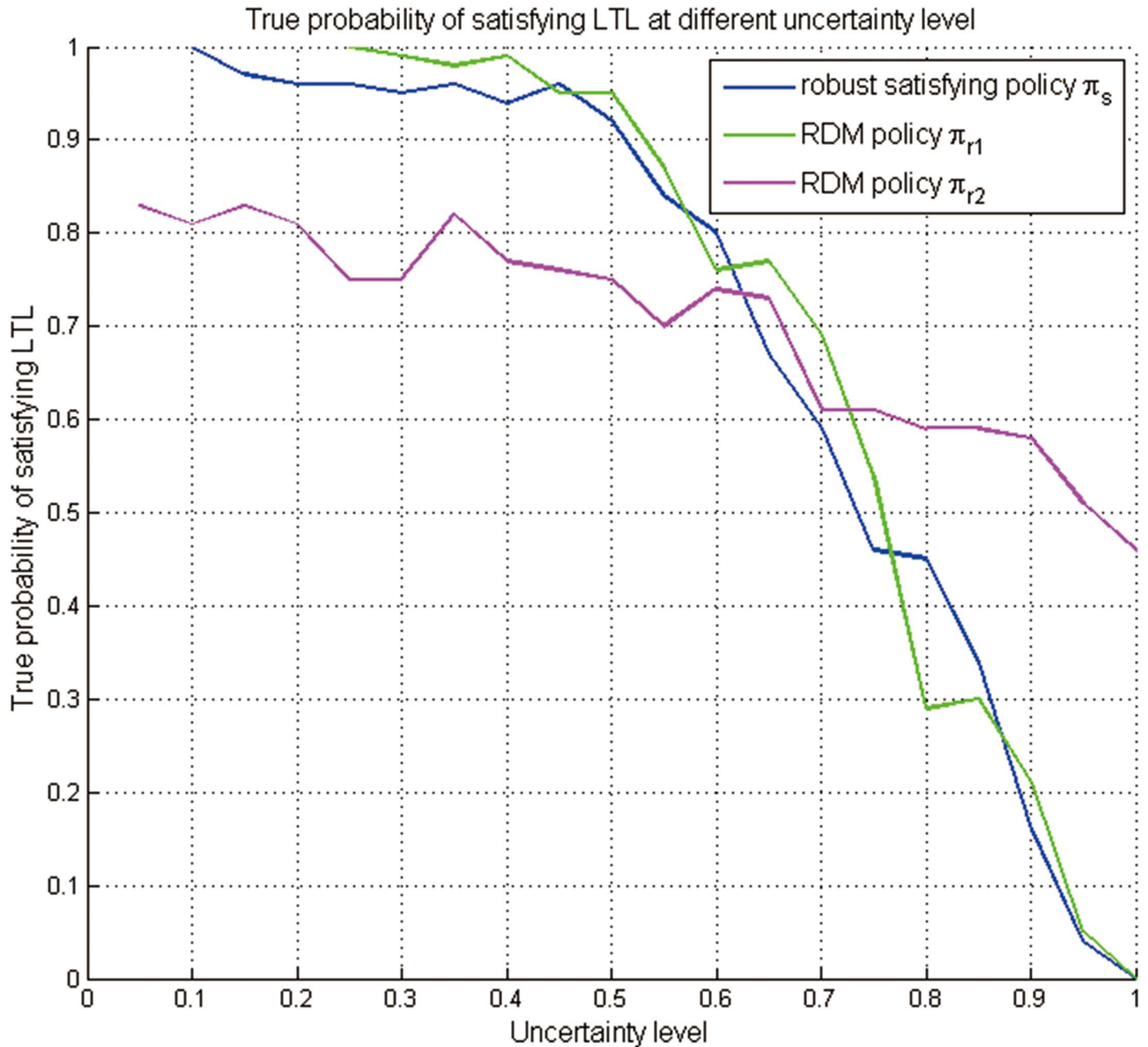
**Table 3. The true LSPs for the min-max robust policy  $\pi_m$  and the robust satisficing policy  $\pi_s$  at different uncertainty levels.**

uncertainty level	0.00	0.10	0.20	0.30	0.40	0.50
LSP for $\pi_m$	1	0.97	0.97	0.86	0.80	0.59
LSP for $\pi_s$	1	1	0.96	0.95	0.94	0.92
uncertainty level	0.60	0.70	0.80	0.90	0.95	1.00
LSP for $\pi_m$	0.48	0.43	0.14	0.02	0.00	0
LSP for $\pi_s$	0.80	0.59	0.45	0.16	0.04	0

doi:10.1371/journal.pone.0166448.t003







**Fig 17. The curves representing the true LSPs for the robust satisficing policy  $\pi_s$  and the RDM robust policies  $\pi_{r1}$  and  $\pi_{r2}$  at different uncertainty levels.**

doi:10.1371/journal.pone.0166448.g017

know or agree on the probability distribution of the key parameters of the model, which can be addressed by the info-gap decision theory; complex missions with rigorous temporal constraints, which can be described by LTL. Potential applications for our method include persistent surveillance task [42], cargo transportation in rough terrain [12], navigation in dangerous environments, and autonomous operation of robots in chaotic work and home environments.

In the future, our method can be extended to many fields of applications. In the field of autonomous robots, we will extend our method to address the multi-robot task cooperation



**Table 4. The true LSPs for the robust satisficing policy  $\pi_s$  and the RDM robust policies  $\pi_{r1}$  and  $\pi_{r2}$  at different uncertainty levels.**

uncertainty level	0.00	0.10	0.20	0.30	0.40	0.50
LSP for $\pi_s$	1	1	0.96	0.95	0.94	0.92
LSP for $\pi_{r1}$	1	1	1	0.99	0.99	0.95
LSP for $\pi_{r2}$	0.85	0.81	0.81	0.75	0.77	0.75
uncertainty level	0.60	0.70	0.80	0.90	0.95	1.00
LSP for $\pi_s$	0.80	0.59	0.45	0.16	0.04	0
LSP for $\pi_{r1}$	0.76	0.69	0.29	0.21	0.05	0
LSP for $\pi_{r2}$	0.70	0.61	0.59	0.58	0.51	0.46

doi:10.1371/journal.pone.0166448.t004

problem with multiple tasks, by introducing the timed automaton and the coordination mechanism for multiple tasks. Besides, the robust satisficing decision making method can be creatively applied to model the severe uncertainty in the economic and climate fields, which can help to generate some robust satisficing economic or climate policies to avoid significant failures.

### Supporting Information

**S1 Appendix. Inner minimization problem.** For solving the inner minimization problem, the dual linear programming is used.

(PDF)

**S1 Data. MDP data.**

(XLSX)

**S2 Data. DRA data.** The LTL formula in the experiment is transformed into a DRA.

(XLSX)

**S3 Data. Generated policies data.** This data set includes the robust satisficing policies  $\pi_1$ ,  $\pi_2$ ,  $\pi_3$ ,  $\pi_4$ , and  $\pi_s$ , the min-max robust policy  $\pi_m$ , and the RDM robust policies  $\pi_{r1}$  and  $\pi_{r2}$ .

(XLSX)

### Acknowledgments

This work was supported by National Natural Science Foundation (NNSF) of China under Grant 61403410.

### Author Contributions

**Conceptualization:** XJ.

**Data curation:** XJ.

**Formal analysis:** XJ YN.

**Investigation:** YN.

**Funding acquisition:** YN.

**Methodology:** XJ.

**Project administration:** LS.

**Resources:** LS.

**Software:** XJ YN.

**Supervision:** YN LS.

**Validation:** YN LS.

**Visualization:** XJ.

**Writing – original draft:** XJ.

**Writing – review & editing:** XJ YN.

## References

1. Goodrich M, Morse B, Gerhardt D. Supporting wilderness search and rescue using a camera-equipped mini UAV: Research Articles. *Journal of Field Robotics*. 2008; 25(1–2):89–110. doi: [10.1002/rob.20226](https://doi.org/10.1002/rob.20226)
2. Ji X, Wang X, Niu Y, Shen L. Cooperative search by multiple unmanned aerial vehicles in a nonconvex environment. *Mathematical Problems in Engineering*; vol. 5; 2015.
3. Semsch E, Jakob M, Pavlicek D, Pechoucek M. Autonomous UAV Surveillance in Complex Urban Environments. In: *IEEE/WIC/ACM International Joint Conferences on Web Intelligence and Intelligent Agent Technologies*; 2009. pp. 82–85.
4. Bulatov D, Solbrig P, Gross H, Wernerus P, Repasi E, Heipke C. Context-Based Urban Terrain Reconstruction from Uav-Videos for Geoinformation Applications. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*. 2012; 3822:75–80. doi: [10.5194/isprsarchives-XXXVIII-1-C22-75-2011](https://doi.org/10.5194/isprsarchives-XXXVIII-1-C22-75-2011)
5. Berman E, Fladeland M, Kolyer R, Gupta M, Liem J. Greenhouse gas analyzer for measurements of carbon dioxide, methane, and water vapor aboard an unmanned aerial vehicle. *Sensors and Actuators B Chemical*. 2012; 169(4):128–135. doi: [10.1016/j.snb.2012.04.036](https://doi.org/10.1016/j.snb.2012.04.036)
6. Iscold P, Pereira G, Torres L. Development of a Hand-Launched Small UAV for Ground Reconnaissance. *IEEE Transactions on Aerospace and Electronic Systems*. 2010; 46(1):335–348. doi: [10.1109/TAES.2010.5417166](https://doi.org/10.1109/TAES.2010.5417166)
7. Bhatia A, Maly MR, Kavraki LE, Vardi MY. Motion planning with complex goals. *IEEE Robotics and Automation Magazine*. 2011; 18(3):55–64. doi: [10.1109/MRA.2011.942115](https://doi.org/10.1109/MRA.2011.942115)
8. Loizou SG, Kyriakopoulos KJ. Automatic synthesis of multi-agent motion tasks based on LTL specifications. In: *IEEE Conference on Decision and Control*. vol. 1; 2005. pp. 153–158.
9. Wongpiromsarn T, Topcu U, Murray RM. Receding Horizon Control for Temporal Logic Specifications. In: *ACM International Conference on Hybrid Systems: Computation & Control*; 2010. pp. 101–110.
10. Puterman ML. *Markov decision processes: Discrete stochastic dynamic programming*. Wiley. 1994; 37(3):353. doi: [10.1002/9780470316887](https://doi.org/10.1002/9780470316887)
11. Xu CD, Smith SL, Belta C, Rus D. LTL Control in Uncertain Environments with Probabilistic Satisfaction Guarantees. *World Congress*. 2011; pp. 3515–3520.
12. Ding X, Smith SL, Belta C, Rus D. Optimal Control of Markov Decision Processes With Linear Temporal Logic Constraints. *IEEE Transactions on Automatic Control*. 2014; 59(5):1244–1257. doi: [10.1109/TAC.2014.2298143](https://doi.org/10.1109/TAC.2014.2298143)
13. Luna R, Lahijanian M, Moll M, Kavraki LE. Asymptotically Optimal Stochastic Motion Planning with Temporal Goals. In: *The Workshop on the Algorithmic Foundations of Robotics*; 2014. pp. 335–352.
14. Baier Christel, Katoen JoostPieter. *Principles of Model Checking*. Mit Press. 2008;.
15. Kupferman O, Vardi MY. *Model Checking of Safety Properties*. *Formal Methods in System Design*. 1998; 19(3):685–685.
16. Lacerda B, Parker D, Hawes N. Optimal and dynamic planning for Markov decision processes with co-safe LTL specifications. In: *IEEE/RSJ International Conference on Intelligent Robots and Systems*; 2014. pp. 1511–1516.
17. Goldfarb D, Iyengar G. Robust Portfolio Selection Problems. *Mathematics of Operations Research*. 2003; 28(1):1–38. doi: [10.1287/moor.28.1.1.14260](https://doi.org/10.1287/moor.28.1.1.14260)
18. Iii CCW, Eldeib HK. Markov decision processes with imprecise transition probabilities. *Operations Research*. 1994; 42(4):739–749. doi: [10.1287/opre.42.4.739](https://doi.org/10.1287/opre.42.4.739)
19. Satia JK, Lave RE. Markovian Decision Processes with Uncertain Transition Probabilities. *Operations Research*. 1973; 21(3):728–740. doi: [10.1287/opre.21.3.728](https://doi.org/10.1287/opre.21.3.728)

20. Goode A. Solving Uncertain Markov Decision Problems. Carnegie Mellon University. 2001.
21. Givan R, Leach S, Dean T. Bounded parameter Markov decision processes. Springer Berlin Heidelberg; 1997. doi: [10.1007/3-540-63912-8\\_89](https://doi.org/10.1007/3-540-63912-8_89)
22. Nilim A, El Ghaoui L. Robust Control of Markov Decision Processes with Uncertain Transition Matrices. *Operations Research*. 2005; 53(5):780–798. doi: [10.1287/opre.1050.0216](https://doi.org/10.1287/opre.1050.0216)
23. Kalyanasundaram S, Chong EKP, Shroff NB. Markov decision processes with uncertain transition rates: sensitivity and robust control. In: *IEEE Conference on Decision and Control*; 2002. pp. 3799–3804 vol.4.
24. Chang HS, Lee HG, Fu MC, Marcus SI. Evolutionary policy iteration for solving Markov decision processes. *IEEE Transactions on Automatic Control*. 2005; 50(11):1804–1808. doi: [10.1109/TAC.2005.858644](https://doi.org/10.1109/TAC.2005.858644)
25. Iyengar GN. Robust Dynamic Programming. *Mathematics of Operations Research*. 2004; 30(2): 257–280. doi: [10.1287/moor.1040.0129](https://doi.org/10.1287/moor.1040.0129)
26. Bertuccelli LF, Wu A, How JP. Robust Adaptive Markov Decision Processes: Planning with Model Uncertainty. *IEEE Control Systems*. 2012; 32(32):96–109. doi: [10.1109/MCS.2012.2205478](https://doi.org/10.1109/MCS.2012.2205478)
27. Wolff EM, Topcu U, Murray RM. Robust control of uncertain Markov Decision Processes with temporal logic specifications. In: *IEEE Conference on Decision and Control*; 2012. pp. 3372–3379.
28. Brown CB. Info-Gap Decision Theory. *Decisions Under Severe Uncertainty*, second ed., Yakov Ben-Haim, Academic Press (2006). *Structural Safety*. 2011;33(1):122.
29. Hu X, Chen Y, Luo H. Robust decision making for UAV air-to-ground attack under severe uncertainty. *Journal of Central South University*. 2015; 22(11):4263–4273. doi: [10.1007/s11771-015-2975-y](https://doi.org/10.1007/s11771-015-2975-y)
30. Manktelow KI. Reasoning and thinking. Psychology Press; 1999.
31. Simon HA. Rational choice and the structure of the environment. *Psychological Review*. 1956; 63(2): 129–138. doi: [10.1037/h0042769](https://doi.org/10.1037/h0042769) PMID: [13310708](https://pubmed.ncbi.nlm.nih.gov/13310708/)
32. Lempert RJ, Groves DG, Popper SW, Bankes SC. A General, Analytic Method for Generating Robust Strategies and Narrative Scenarios. *Management Science*. 2006; 52(4):514–528. doi: [10.1287/mnsc.1050.0472](https://doi.org/10.1287/mnsc.1050.0472)
33. Regan HM, Ben-Haim Y, Langford B, Wilson WG, Lundberg P, Andelman SJ, et al. Robust Decision Making under Severe Uncertainty for Conservation Management. *Ecological Applications*. 2005; 15(4): 1471–1477. doi: [10.1890/03-5419](https://doi.org/10.1890/03-5419)
34. P S Gareth, BH Yakov, W Keith, M Graeme. Evaluation of neural network robust reliability using information-gap theory. *IEEE Transactions on Neural Networks*. 2006; 17(6):1349–61. doi: [10.1109/TNN.2006.880363](https://doi.org/10.1109/TNN.2006.880363) PMID: [17131652](https://pubmed.ncbi.nlm.nih.gov/17131652/)
35. Itay S, Tal S, Yakov BH. Info-Gap Approach to Multiagent Search Under Severe Uncertainty. *IEEE Transactions on Robotics*. 2011; 26(6):1032–1041.
36. Mascareñas D, Farrar C. Development of an info-gap-based path planner to enable nondeterministic low-observability mobile sensor nodes. In: *Society of Photo-Optical Instrumentation Engineers Conference*. vol. 26; 2012. pp. 8387191–83871910.
37. Lempert R, Popper S, Bankes S. *Shaping the Next One Hundred Years: New Methods for Quantitative, Long-Term Policy Analysis*. Santa Monica, CA: RAND Corporation, MR-1626-RPC, 2003.
38. Kwiatkowska M, Norman G, Parker D. PRISM 4.0: Verification of Probabilistic Real-Time Systems. *Lecture Notes in Computer Science*. 2010; 6806:585–591. doi: [10.1007/978-3-642-22110-1\\_47](https://doi.org/10.1007/978-3-642-22110-1_47)
39. Klein J, Baier C. Experiments with deterministic  $\omega$ -automata for formulas of linear temporal logic. *Theoretical Computer Science*. 2005; 363(2):182–195.
40. Mcinerney D, Lempert R, Keller K What are robust strategies in the face of uncertain climate threshold responses? *Climatic Change*. 2009; 112(3):547–568.
41. Lempert R, Collins M. Managing the Risk of Uncertain Threshold Responses: Comparison of Robust, Optimum, and Precautionary Approaches. *Risk Analysis*. 2007; 27(4):1009–26. doi: [10.1111/j.1539-6924.2007.00940.x](https://doi.org/10.1111/j.1539-6924.2007.00940.x) PMID: [17958508](https://pubmed.ncbi.nlm.nih.gov/17958508/)
42. Ulusoy A, Belta C. Receding horizon temporal logic control in dynamic environments. *International Journal of Robotics Research*. 2014; 33(12):1593–1607. doi: [10.1177/0278364914537008](https://doi.org/10.1177/0278364914537008)