

Thoughts on Writing a Good (Robotics) Paper

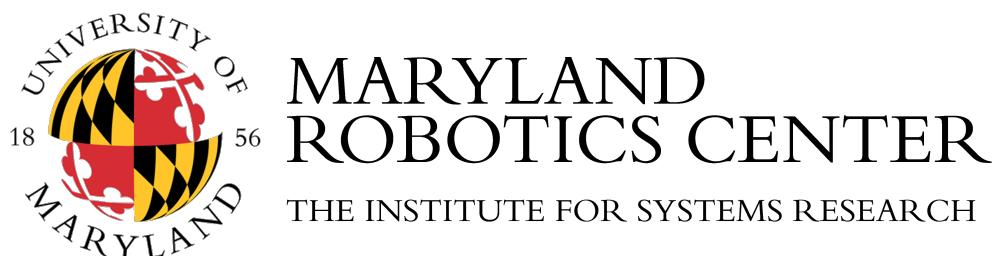
Pratap Tokekar

Assistant Professor

Computer Science

University of Maryland

<http://tokekar.com/>



- Disclaimer: Robotics is an interdisciplinary and broad field. What I'm going to present is my personal view of the publishing landscape in robotics.

Agenda

- ▶ Writing a good paper
- ▶ Robotics conference/journals landscape

Writing a good paper

- ▶ How to craft “the research story” for a paper?
- ▶ How to convey the fundamental science or technology impact of your work?

Course Project → Research Paper

- ▶ Course project
 - ▶ Advances the knowledge of the students
 - ▶ Students learn something new that they didn't know
- ▶ Research paper
 - ▶ Advances the knowledge of the community
 - ▶ Researchers learn something new that the community didn't know
- ▶ Goal of a research paper is to disseminate the new knowledge to the rest of the community

Your paper needs to

1. What did the community know before you did whatever you did?
2. What are the new things you learned after you did whatever you did?
3. What exactly did you do?

Your paper needs to

1. What did the community know before you did whatever you did?
2. What are the new things you learned after you did whatever you did?
3. What exactly did you do?

A common rookie mistake is to focus only on the “what did you do?” question and ignore the first two. Difference between a course project report and a research paper.

Your paper needs to

1. What did the community know before you did whatever you did?
2. What are the new things you learned after you did whatever you did?
3. What exactly did you do?

If you can't answer the questions, then something is wrong... need to take a step back.

Your paper needs to

1. What did the community know before you did whatever you did?
2. What are the new things you learned after you did whatever you did?
3. What exactly did you do?

You should constantly be asking yourself the first two questions. Even before you start working on the problem (hypothesize the answers). The answers may change as your research evolves -- that's okay.

- **Introduction**
 - Overview of Q1, Q2, Q3; plus
 - *Why should the community care?*
- **Related Work**
 - Q1
- **Problem Formulation**
- **Algorithm/Methodology**
- **Evaluation**
 - Q2 & Q3
- **Conclusion and Future Work**
 - Overview of Q1, Q2, and Q3; plus
 - *What does the community still not know?*

Your paper needs to

1. What did the community know before you did whatever you did?
2. What are the new things you learned after you did whatever you did?
3. What exactly did you do?
4. Why should the community care?
5. What does the community still not know?

An Approximation Algorithm for Risk-averse Submodular Optimization

Lifeng Zhou and Pratap Tokekar

Department of Electrical & Computer Engineering, Virginia Tech, USA.
`{lfzhou, tokekar}@vt.edu`

Abstract. We study the problem of incorporating risk while making combinatorial decisions under uncertainty. We formulate a discrete submodular maximization problem for selecting a set using Conditional-Value-at-Risk (CVaR), a risk metric commonly used in financial analysis. While CVaR has recently been used in optimization of linear cost functions in robotics, we take the first stages towards extending this to discrete submodular optimization and provide several positive results. Specifically, we propose the Sequential Greedy Algorithm that provides an approximation guarantee on finding the maxima of the CVaR cost function under a matroidal constraint. The approximation guarantee shows that the solution produced by our algorithm is within a constant factor of the optimal and an additive term that depends on the optimal. Our analysis uses the curvature of the submodular set function, and proves that the algorithm runs in polynomial time. This formulates a number of combinatorial optimization problems that appear in robotics. We use two such problems, vehicle assignment under uncertainty for mobility-on-demand and sensor selection with failures for environmental monitoring, as case studies to demonstrate the efficacy of our formulation.

1 Introduction

Combinatorial optimization problems find a variety of applications in robotics. Typical examples include:

- *Sensor placement*: Where to place sensors to maximally cover the environment [1] or reduce the uncertainty in the environment [2]?
- *Task allocation*: How to allocate tasks to robots to maximize the overall utility gained by the robots [3]?
- *Combinatorial auction*: How to choose a combination of items for each player to maximize the total rewards [4]?

Algorithms for solving such problems find use in sensor placement for environment monitoring [1, 2], robot-target assignment and tracking [5–7], and informative path planning [8]. The underlying optimization problem in most cases can be written as:

$$\max_{S \in \mathcal{I}, S \subseteq \mathcal{X}} f(S), \quad (1)$$

Q4. Why should the community care?

where \mathcal{X} denotes a ground set from which a subset of elements S must be chosen. f is a monotone submodular utility function [9, 10]. Submodularity is the property of diminishing returns. Many information theoretic measures, such as mutual information [2], and geometric measures such as the visible area [11], are known to be submodular. \mathcal{I} denotes a matroidal constraint [9, 10]. Matroids are a powerful combinatorial tool that can represent constraints on the solution set, e.g., cardinality constraints (“place no more than k sensors”) and connectivity constraints (“the communication graph of the robots must be connected”) [12]. The objective of this problem is to find a set \mathcal{S} satisfying a matroidal constraint \mathcal{I} and maximizing the utility $f(\mathcal{S})$. The general form of this problem is NP-complete. However, a greedy algorithm yields a constant factor approximation guarantee [9, 10].

In practice, sensors can fail or get compromised [13] or robots may not know the exact positions of the targets [14]. Hence, the utility $f(\mathcal{S})$ is not necessarily deterministic but can have uncertainty. Our main contribution is to extend the traditional formulation given in Eq. 1 to also account for the uncertainty in the actual cost function. We model the uncertainty by assuming that the utility function is of the form $f(\mathcal{S}, y)$ where $\mathcal{S} \in \mathcal{X}$ is the decision variable and $y \in \mathcal{Y}$ represents a random variable which is independent of \mathcal{S} . We focus on the case where $f(\mathcal{S}, y)$ is monotone submodular in $\mathcal{S} \in \mathcal{X}$ and integrable in y .

The traditional way of stochastic optimization is to use the expected utility as the objective function: $\max_{\mathcal{S} \in \mathcal{I}, \mathcal{S} \in \mathcal{X}} \mathbb{E}_y[f(\mathcal{S}, y)]$. Since the sum of the monotone submodular functions is monotone submodular, $\mathbb{E}_y[f(\mathcal{S}, y)]$ is still monotone submodular in \mathcal{S} . Thus, the greedy algorithm still retains its constant-factor performance guarantee [9, 10]. Examples of this approach include influence maximization [15], moving target detection and tracking [14], and robot assignment with travel-time uncertainty [16].

While optimizing the expected utility has its uses, it also has its pitfalls. Consider the example of mobility-on-demand where two self-driving vehicles, v_1 and v_2 , are available to pick up the passengers at a demand location (Fig. 1). v_1 is closer to the demand location, but it needs to cross an intersection where it may need to stop and wait. v_2 is further from the demand location but there is no intersection along the path. The travel time for v_1 follows a bimodal distribution (with and without traffic stop) whereas that for v_2 follows a unimodal distribution with a higher mean but lower uncertainty. Clearly, if the passenger uses the expected travel

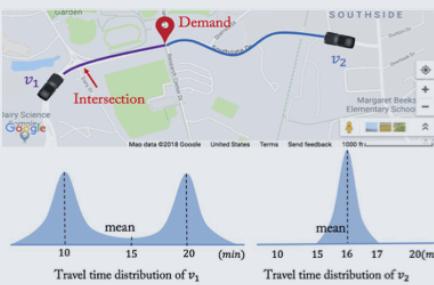


Fig. 1. Mobility on demand with travel time uncertainty of self-driving vehicles.

Q4. Why should the community care?

Q1. What did the community know before you did whatever you did?

time as the objective, they would choose v_1 . However, they will risk waiting a much longer time, i.e., $17 \sim 20$ min about half of the times. A more risk-averse passenger would choose v_2 which has higher expected waiting time 16 min but a lesser risk of waiting longer.

Thus, in these scenarios, it is natural to go beyond expectation and focus on a risk-averse measure. One popular coherent risk measure is *Conditional-Value-at-Risk* (CVaR) [17, 18]. CVaR takes a risk level α which is the probability of the worst α -tail cases. Loosely speaking, maximizing CVaR is equivalent to maximizing the expectation of the worst α -tail scenarios.¹ This risk-averse decision is rational especially when the failures can lead to unrecoverable consequences, such as a sensor failure.

Related work. Yang and Chakraborty studied a chance-constrained combinatorial optimization problem that takes into account the risk in multi-robot assignment [19]. They later extended this to knapsack problems [20]. They solved the problem by transforming it to a risk-averse problem with mean-variance measure [21]. Chance-constrained optimization is similar to optimizing the Value-at-Risk (VaR), which is another popular risk measure in finance [22]. However, Majumdar and Pavone argued that CVaR is a better measure to quantify risk than VaR or mean-variance based on six proposed axioms in the context of robotics [23].

Several works have focused on optimizing CVaR. In their seminal work [18], Rockafellar and Uryasev presented an algorithm for CVaR minimization for reducing the risk in financial *portfolio optimization* with a large number of instruments. Note that, in portfolio optimization, we select a distribution over available decision variables, instead of selecting a single one. Later, they showed the advantage of optimizing CVaR for general loss distributions in finance [24].

When the utility is a discrete submodular set function, i.e., $f(\mathcal{S}, y)$, Maehara presented a negative result for maximizing CVaR [25]—there is no polynomial time multiplicative approximation algorithm for this problem under some reasonable assumptions in computational complexity. To avoid this difficulty, Ohsaka and Yoshida in [26] used the same idea from portfolio optimization and proposed a method of selecting a distribution over available sets rather than selecting a single set, and gave a provable guarantee. Following this line, Wilder considered a CVaR maximization of a continuous submodular function instead of the submodular set functions [27]. They gave a $(1 - 1/e)$ -approximation algorithm for continuous submodular functions and also evaluated the algorithm for discrete submodular functions using portfolio optimization [26].

Contributions. We focus on the problem of selecting a single set, similar to [25], to maximize CVaR rather than portfolio optimization [26, 27]. This is because we are motivated by applications where a one-shot decision (placing sensors and assigning vehicles) must be taken. Our contributions are as follows:

- We propose the Sequential Greedy Algorithm (SGA) which uses the deterministic greedy algorithm [9, 10] as a subroutine to find the maximum value of CVaR (Algorithm 1).

¹ We formally review CVaR and other related concepts in Section 2.1

Q1. What did the community know before you did whatever you did?

Q3. Why exactly did you do?

- We prove that the solution found by SGA is within a constant factor of the optimal performance along with an additive term which depends on the optimal value. We also prove that SGA runs in polynomial time (Theorem 1) and the performance improves as the running time increases.
- We demonstrate the utility of the proposed CVaR maximization problem through two case studies (Section 3.2). We evaluate the performance of SGA through simulations (Section 5).

Organization of rest of the paper. We give the necessary background knowledge for the rest of the paper in Section 2. We formulate the CVaR submodular maximization problem with two case studies in Section 3. We present SGA along with the analysis of its computational complexity and approximation ratio in Section 4. We illustrate the performance of SGA to the two case studies in Section 5. We conclude the paper in Section 6.

2 Background and Preliminaries

We start by defining the conventions used in the paper.

Calligraphic font denotes a set (e.g., \mathcal{A}). Given a set \mathcal{A} , $2^{\mathcal{A}}$ denotes its power set. $|\mathcal{A}|$ denotes the cardinality of \mathcal{A} . Given a set \mathcal{B} , $\mathcal{A} \setminus \mathcal{B}$ denotes the set of elements in \mathcal{A} that are not in \mathcal{B} . $\Pr[\cdot]$ denotes the probability of an event and $\mathbb{E}[\cdot]$ denotes the expectation of a random variable. $[x] = \min\{n \in \mathbb{Z} | x \leq n\}$ where \mathbb{Z} denotes the set of integers.

Next, we give the background on set functions (in the appendix file) and risk measures.

2.1 Risk measures

Let $f(\mathcal{S}, y)$ be a utility function with decision set \mathcal{S} and the random variable y . For each \mathcal{S} , the utility $f(\mathcal{S}, y)$ is also a random variable with a distribution induced by that of y . First, we define the Value-at-Risk at risk level $\alpha \in (0, 1]$.

Value at Risk:

$$\text{VaR}_\alpha(\mathcal{S}) = \inf\{\tau \in \mathbb{R}, \Pr[f(\mathcal{S}, y) \leq \tau] \geq \alpha\}. \quad (2)$$

Thus, $\text{VaR}_\alpha(\mathcal{S})$ denotes the left endpoint of the α -quantile(s) of the random variable $f(\mathcal{S}, y)$. The Conditional-Value-at-Risk is the expectation of this set of α -worst cases of $f(\mathcal{S}, y)$, defined as:

Q2. Why does the community know after you did whatever you did?

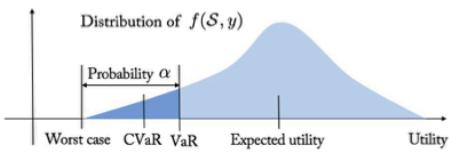


Fig. 2. An illustration of risk measures: VaR and CVaR.

How to get a paper rejected?

- Describe only what you did but not what you learned
 - *Reviewers can't answer Q2. Reject.*
- Focus only on the technical ideas but miss the big picture
 - *Reviewers can't answer Q4. Reject.*
- Write a paper without reading other papers
 - *Can't answer Q1 & Q2. Reject.*
- Use vague language and weasel words
 - *Can't answer Q3 & Q5. Reject.*

- How to write a good (robotics) conference paper?
- How to convert a conference paper into a good journal paper?

3 Step Strategy

1. Tell them what you are going to tell them
 2. Then actually tell them
 3. Then tell them what you just told them
-
- This 3 step strategy works at every level
 - Paper
 - Section
 - Sub-section
 - Paragraph

Start with an outline of the paper

- Introduction
- Related Work
- Problem Formulation
- Algorithm/Methodology
- Evaluation (with Discussion)
- Conclusion and Future Work

Then a slightly more detailed outline

- Introduction
 - Q4
 - Q1, Q2, Q3
 - List of contributions
- Related Work
 - Q2 for area #1
 - Q2 for area #2
- Problem Formulation
- Algorithm/Methodology
 - Version 1
 - Version 2
- Evaluation
 - Simulations
 - Experiments
 - Discussion (Q2)
- Conclusion and Future Work
 - Q3
 - Q5

Then an even more detailed outline

- Remember the 3 step strategy
- Start writing down what you will be saying in each subsection, in each paragraph..
- Filling in the details after that is easy
- Iterate, iterate, iterate
- Spend more time editing than writing!

Notation and Problem Formulation

- You should have a separate section / sub-section where you
 - describe the formal notation in your paper
 - give detailed list of assumptions
 - give precise definitions of terms that you are using
 - give precise, formal definition of the problem
- Q3: If the reviewer is unsure about what you did, then they may make the most uncharitable interpretation.

3 Problem Formulation

In this section, we formally define the terminology and the two problems studied. We assume that the environment is a compact domain $U \subset \mathbb{R}^2$. We make the following assumptions about the spatial function $f_x \equiv f(x)$:

Assumption 1 (Smoothness) *The true function is smooth in the sense of Lipschitz [21], i.e., $\forall x_i, x_j \in U : |f_{x_i} - f_{x_j}| \leq \mathcal{L} \|x_i - x_j\|_2$, where \mathcal{L} is the Lipschitz smoothness constant.*

Assumption 2 (Kernel) *The GP regression uses a squared-exponential kernel [7]. The hyperparameters of the kernel (length scale and signal variance) are known a priori.*

Assumption 3 (Bound on Measurements) *Optimal algorithms for both problems satisfy the chance constraints using a finite number of measurements, no more than N .*

Let X denote the set of measurement locations within U produced by an algorithm. In the placement problem, our goal is to minimize the cardinality of X whereas in the path planning problem the goal is to minimize the time required to visit and obtain measurements at X . The two problems are formally defined below.

Problem 1 (Placement). Given Assumptions 1–3, find the minimum number of measurement locations in U , such that the posterior GP prediction for any point in U is within $\pm\epsilon$ of the actual value with probability at least δ , *i.e.*,

$$\begin{aligned} & \text{minimize} && \text{number of measurement locations } |X| \\ & \text{subject to} && \Pr\{|\hat{\mu}(x) - f(x)| \leq \epsilon\} \geq \delta, \forall x \in U \end{aligned}$$

where $f(x)$ is the actual function value that is to be estimated at a point $x \in U$, and $\hat{\mu}(x)$ is the predicted value using measurements obtained at locations X .

Problem 2 (Mobile). Given Assumptions 1–3, find a minimum time trajectory for a mobile sensor that obtains a finite set of measurements at one or more locations in U , such that the posterior GP prediction for any point in U is within $\pm\epsilon$ of the true value with probability at least δ , *i.e.*,

$$\begin{aligned} & \text{minimize} && \text{len}(\tau) + \eta\beta(X), \\ & \text{subject to} && \Pr\{|\hat{\mu}(x) - f(x)| \leq \epsilon\} \geq \delta, \forall x \in U. \end{aligned}$$

τ denotes the tour of the robot. Assume that the robot travels at unit speed, obtains one measurement in η units of time and obtains $\beta(X)$ total measurements at the measurement locations X . Here, we use the function $\beta(X)$ to take into account the fact that the robot may obtain more than one measurement at a specific location. Therefore, the number of measurements can be more than $|X|$.

Not just for “theoretical” papers

- Even if you are writing an experimental paper, you should be precise in defining the problem
- Think of it this way: your problem formulation section defines the scope of your work
 - What are the inputs?
 - What are the assumptions / restrictions on the inputs?
 - What are the expected outputs?
 - etc.

The input to our algorithm is a set of n sites, x_i , that must be visited by the UAV. We start with a list of common assumptions: (1) unit rate of discharge (1% per second); (2) UAV has an initial battery charge of 100%; (3) UAV and UGVs start at a common *depot*, d ; (4) all the sites are at the same altitude; (5) UAV can fly between any two sites if it starts at 100% battery level; (6) UGVs have unlimited fuel/battery capacity. All but the last assumption are only for the sake of convenience and ease of presentation and can be easily relaxed. Although UGVs cannot have unlimited operational time, it is a reasonable assumption since UGVs can have much larger batteries or can be refueled quickly.

Robotics Conferences

ICRA 2020

31 May • 4 June 2020
Palais des Congrès de Paris
Paris • FRANCE



ROBOTICS
SCIENCE AND SYSTEMS

Deadline: September 15th

Deadline: January 30th

2020 IEEE/RSJ

International Conference on
Intelligent Robots and Systems(IROS)

October 25-29, 2020 Las Vegas, NV, USA

Deadline: March 1st

Robotics Conferences

ICRA 2020

31 May • 4 June 2020
Palais des Congrès de Paris
Paris • FRANCE



ROBOTICS
SCIENCE AND SYSTEMS

Deadline: September 15th

Deadline: January 30th

Deadline: March 1st

- ▶ DARS (every even year)
- ▶ MRS (every odd year)
- ▶ WAFR (every even year)
- ▶ ISER (every even year)
- ▶ CoRL (every year)
- ▶ ISRR (every odd year)

distributed robotics
multi-robot systems
algorithmic robotics
experimental robotics
robot learning
position papers

2020 IEEE/RSJ

International Conference on
Intelligent Robots and Systems(IROS)

October 25-29, 2020 Las Vegas, NV, USA

Journals

- ▶ IEEE TRO *6.483**
- ▶ IJRR *6.134*
- ▶ IEEE TASE *5.224*
- ▶ JFR *4.345* *field robotics*
- ▶ IEEE RAM *4.25* *magazine/popular articles*
- ▶ AuRO *3.634*
- ▶ *IEEE RA-L* *not yet* *short papers, fast decision*

*Impact factors. Frankly, ignore them. All of these venues are good.

Conference vs. Journal

- 6 page conference paper published at ICRA / IROS
 - Demonstrates that the idea *is sound and promising*
 - Gather feedback from audience

evolves into

- 10+ page journal paper
 - *Comprehensively evaluated*
 - ~40% addition to the conference paper

ICRA/IROS + RAL



ICRA+RAL: Sep 10th
ICRA: Sep 15th

2020 IEEE/RSJ

International Conference on
Intelligent Robots and Systems(IROS)

October 25-29, 2020 Las Vegas, NV, USA

IROS+RAL: Feb 24th
IROS: March 1st

Papers can be submitted to ICRA and IROS through two ways:

1. Conference only: up to 6 pages; can be later submitted to a journal with suitable additions
2. Conference+RAL: up to 8 pages; independently reviewed

ICRA only or ICRA+RAL

- If the work is comprehensively evaluated and can fit in 8 pages → submit to ICRA+RAL
- If the idea is promising and sound but has potential to be more thoroughly evaluated →
 - submit to ICRA only
 - then expand the work and submit to other journals

- ▶ ICRA: *accept* RAL: *accept*
 - ▶ Present the paper at ICRA
 - ▶ Published in RAL but not in conference proceedings
 - ▶ Cannot submit it again to another conference/journal
- ▶ ICRA: *accept* RAL: *reject*
 - ▶ Published in the ICRA proceedings
 - ▶ Present the paper at ICRA
 - ▶ Can submit *an expanded version* to another journal
- ▶ ICRA: *reject* RAL: *accept*
 - ▶ Published in RAL
 - ▶ Cannot present at ICRA
 - ▶ Cannot submit it to another conference/journal
- ▶ ICRA: *reject* RAL: *reject*
 - ▶ Sorry!

- ▶ ICRA: *accept* RAL: *accept*
 - ▶ Present the paper at ICRA
 - ▶ Published in RAL but not in conference proceedings
 - ▶ Cannot submit it again to another conference/journal
- ▶ ICRA: *accept* RAL: *reject*
 - ▶ Published in the ICRA proceedings
 - ▶ Present the paper at ICRA
 - ▶ Can submit *an expanded version* to another journal
- ▶ ICRA: *reject* RAL: *accept*
 - ▶ Published in RAL
 - ▶ Cannot present at ICRA
 - ▶ Cannot submit it to another conference/journal
- ▶ ICRA: *reject* RAL: *reject*
 - ▶ Sorry!

- ▶ ICRA: *accept* RAL: *accept*
 - ▶ Present the paper at ICRA
 - ▶ Published in RAL but not in conference proceedings
 - ▶ Cannot submit it again to another conference/journal
- ▶ ICRA: *accept* RAL: *reject*
 - ▶ Published in the ICRA proceedings
 - ▶ Present the paper at ICRA
 - ▶ Can submit *an expanded version* to another journal
- ▶ ICRA: *reject* RAL: *accept*
 - ▶ Published in RAL
 - ▶ Cannot present at ICRA
 - ▶ Cannot submit it to another conference/journal
- ▶ ICRA: *reject* RAL: *reject*
 - ▶ Sorry!

- ▶ ICRA: *accept* RAL: *accept*
 - ▶ Present the paper at ICRA
 - ▶ Published in RAL but not in conference proceedings
 - ▶ Cannot submit it again to another conference/journal
- ▶ ICRA: *accept* RAL: *reject*
 - ▶ Published in the ICRA proceedings
 - ▶ Present the paper at ICRA
 - ▶ Can submit *an expanded version* to another journal
- ▶ ICRA: *reject* RAL: *accept*
 - ▶ Published in RAL
 - ▶ Cannot present at ICRA
 - ▶ Cannot submit it to another conference/journal
- ▶ ICRA: *reject* RAL: *reject*
 - ▶ Sorry!

I've had the pleasure of all four outcomes

- ▶ ICRA '19 #1: *accept* RAL '19 #1: *accept*
- ▶ ICRA '19 #2: *accept* RAL '19 #2: *reject*
- ▶ ICRA '19 #3: *reject* RAL '19 #3: *reject*
- ▶ ICRA '17: *reject* RAL '17: *accept*

ICRA decisions are one-shot: cannot respond to reviewers comments.

RAL usually gives an option for one resubmission and responding to reviewer comments.

Misconceptions about robotics papers

Conference papers do not need to include any evaluation since that is the job of a journal paper.

No.

- You still need to demonstrate the idea is sound and promising.
- Hard to do it without some argument (qualitative, quantitative, mathematical). It is a matter of degree.
- Ultimately it is subjective and at the discretion of the reviewers and editors.

Misconceptions about robotics papers

Hardware experiments on actual robots are necessary.
No.

- Since 2018, my group has published exactly 16 papers with hardware/real-data experiments and 16 without any hardware/real-data experiments
- You must still demonstrate that your idea is sound and conduct rigorous evaluation
- You can do that through various means: *proofs, simulations, experiments, datasets*

Misconceptions about robotics papers

Hardware experiments on actual robots are sufficient.
No.

- Difference between proof-of-concept hardware demonstrations and rigorous experimental evaluation.
- The latter may be sufficient but the former is not.
Need to back it up with more simulations / proofs.
- Recall Q3: *what did you learn?*

Misconceptions about robotics papers

My algorithm/controller/system design has to be perfect in all conditions. No.

- A common mistake is to expect perfection from the paper.
- But you must identify exactly what conditions it works in, how well it works, and when it does not work.
- Recall Q3 (*What exactly did you do?*) and Q5 (*What do we still not know?*)

Take-Home Message

- Remember the 5 Q's
- 3 step strategy for writing a paper
- Conference papers must demonstrate the idea is sound and promising
- Journal papers must comprehensively evaluate the idea