Enhancing Robot Navigation Policies with Task-Specific Uncertainty Managements

Extended Abstract

Gokul Puthumanaillam
University of Illinois
Urbana-Champaign
Urbana, IL, USA
gokulp2@illinois.edu

Paulo Padrao
Florida International University
Miami, FL, USA
plope113@fiu.edu

Jose Fuentes
Florida International University
Miami, FL, USA
jfuen099@fiu.edu

Leonardo Bobadilla
Florida International University
Miami, FL, USA
bobadilla@cs.fiu.edu

Melkior Ornik
University of Illinois
Urbana-Champaign
Urbana, IL, USA
mornik@illinois.edu

ABSTRACT

Robots navigating complex environments must manage uncertainty from sensor noise, environmental changes, and incomplete information, with different tasks requiring varying levels of precision in different areas. For example, precise localization may be crucial near obstacles but less critical in open spaces. We present GUIDE (Generalized Uncertainty Integration for Decision-Making and Execution), a framework that integrates these task-specific requirements into navigation policies via Task-Specific Uncertainty Maps (TSUMs). By assigning acceptable uncertainty levels to different locations, TSUMs enable robots to adapt uncertainty management based on context. When combined with reinforcement learning, GUIDE learns policies that balance task completion and uncertainty management without extensive reward engineering. Real-world tests show significant performance gains over methods lacking task-specific uncertainty awareness.

KEYWORDS

Uncertainty-guided planning; Task-specific decision-making; Reinforcement learning; Robot navigation

ACM Reference Format:
Gokul Puthumanaillam, Paulo Padrao, Jose Fuentes, Leonardo Bobadilla, and Melkior Ornik. 2025. Enhancing Robot Navigation Policies with Task-Specific Uncertainty Managements: Extended Abstract. In *Proc. of the 24th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2025), Detroit, Michigan, USA, May 19 – 23, 2025*, IFAAMAS, 3 pages.

1 INTRODUCTION

Robots encounter varying degrees of uncertainty in sensor measurements, motion models, and environmental conditions. Critically, not all tasks require uniform levels of certainty: for instance, a robot navigating tight corridors must localize precisely, while crossing open spaces may tolerate higher uncertainty. This *context-dependent* nature of uncertainty requirements motivates a need for policies that can selectively reduce or accept uncertainty based on task-specific demands.

Existing methods either aim to minimize uncertainty everywhere [8, 9, 22] or enforce fixed thresholds [5, 26], often leading to inefficiency when uncertainty requirements vary within the same task. Some approaches attempt a uniform trade-off between task performance and uncertainty management [24, 30, 33], but they still treat uncertainty needs as globally homogeneous. Moreover, reward engineering or manual tuning [4, 40] is frequently required to capture task constraints. In parallel, extensive work in probabilistic robotics [7, 11, 14, 32] addresses uncertainty but does so uniformly across the workspace. Reinforcement learning (RL) has proven effective for navigation [23, 34, 35] but seldom integrates task-specific uncertainty considerations [6, 19]. Techniques that penalize high-uncertainty actions [10, 17, 29, 39] or use Bayesian RL [1, 2, 15, 37] still adopt a one-size-fits-all approach. Risk-aware planning and risk-sensitive RL [13, 16, 20, 31, 38] similarly rely on uniform thresholds. Although language-conditioned methods [3, 21, 28] broaden a robot's task repertoire, they lack a direct mechanism to represent and leverage region-specific uncertainty limits.

We address this gap by introducing *Task-Specific Uncertainty Maps (TSUMs)* that encode allowable uncertainty levels across different regions of the environment for a given task. TSUM captures the *context-dependent value of certainty*, enabling robots to focus on precision only where it is crucial. We present a policy-conditioning framework, *GUIDE* (Generalized Uncertainty Integration for Decision-Making and Execution), which integrates TSUMs into navigation policies. We adapt Soft Actor-Critic (SAC) to *GUIDEd SAC*, balancing task objectives and uncertainty reduction without ad-hoc reward engineering.

2 METHODOLOGY

Consider a robot operating in a continuous state space S with a continuous action set A. Given a navigation task τ specified in natural language, the objective is to learn a policy $\pi(a \mid s)$ that jointly satisfies task objectives and task-specific uncertainty requirements.

Proc. of the 24th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2025), Y. Vorobeychik, S. Das, A. Nowé (eds.), May 19 – 23, 2025, Detroit, Michigan, USA. © 2025 International Foundation for Autonomous Agents and Multiagent Systems (www.ifaamas.org).

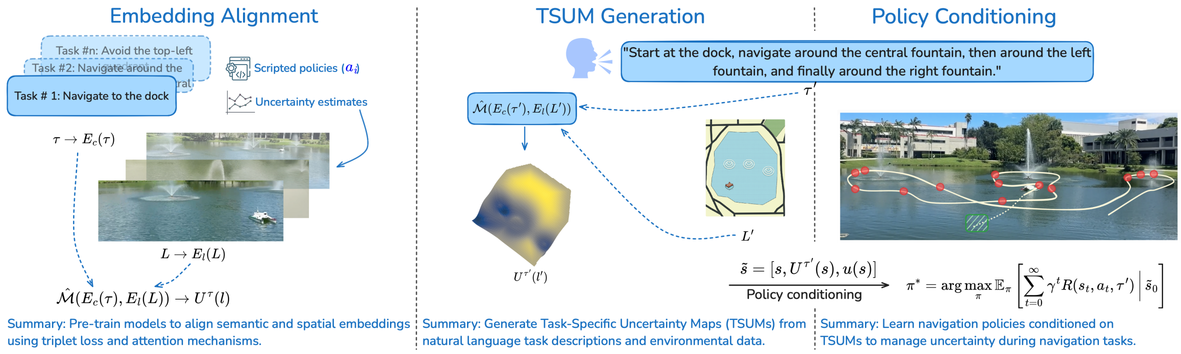

Figure 1: During pretraining, semantic and spatial embeddings are aligned via triplet loss and attention. At deployment, TSUMs derived from task descriptions and environment data condition the navigation policy for task-aware uncertainty management.

Task-Specific Uncertainty Maps (TSUM): A TSUM is defined as $U^\tau(l)$, a scalar representing the acceptable uncertainty at location l. Formally, $U^\tau(l) = w_\Phi\, \Phi^\tau(l) + w_C\, C^\tau(l) + w_\mathcal{E}\, \mathcal{E}(l)$, where $\Phi^\tau(l)$ captures the *relevance* of l for the task, $C^\tau(l)$ encodes *constraints* such as safety and legal restrictions, and $\mathcal{E}(l)$ reflects *environmental factors*. Task semantics are extracted from the natural language specification τ using a RoBERTa-based parser [18], which identifies subtasks and constraints. Each location l is mapped to a spatial embedding via a neural network that processes coordinate and environment features. Alignment between these spatial embeddings and the text-derived subtasks is enforced through a triplet loss that brings related concepts closer in embedding space while separating unrelated pairs. An attention mechanism [36] weights the most relevant subtasks or constraints for each location, producing a single scalar $U^\tau(l)$ indicating how critical it is to maintain low uncertainty at l. Figure 1 (center) represents an example TSUM.

Policy Conditioning: Once the TSUM is generated, the robot's state is augmented with both the TSUM value and the robot's current uncertainty. Concretely, if s denotes the original state, then the augmented state is $\tilde{s} = [s,\ U^\tau(s),\ u(s)]$, where $u(s)$ represents the robot's current state-estimation uncertainty at s. The policy $\pi(a \mid \tilde{s})$ thus explicitly observes how precise the localization at s needs to be, enabling actions that selectively reduce uncertainty in regions with stricter tolerances.

To implement this, we adopt a variant of SAC [12], referred to as GUIDEd SAC (G-SAC). Standard SAC updates the policy and Q-value networks to maximize expected reward while promoting exploratory behavior via an entropy term. In G-SAC, the Q-function and policy networks receive $\tilde{s}$ as input, allowing them to incorporate the TSUM-derived acceptable uncertainty as part of the state. By conditioning on $U^\tau(s)$, the learned policy mitigates uncertainty precisely in those regions where the task demands high localization accuracy, while avoiding unnecessary effort in areas where uncertainty can be higher without affecting task success–eliminating the need for extensive reward tuning or ad hoc penalty terms, as the TSUM itself encodes spatially varying uncertainty requirements.

3 EXPERIMENTS AND RESULTS

We evaluate GUIDE on real-world navigation tasks using an autonomous surface vehicle (ASV) operating in a lake with obstacles and environmental disturbances. By default, the ASV employs noisy state estimation (low cost, high uncertainty) but can temporarily request precise GPS data at an added penalty [25, 27].

Baselines and Ablations. We compare GUIDE (implemented as G-SAC) against: **SAC** (no TSUMs), **SAC-P** (penalized uncertainty), **B-SAC** (bootstrapped uncertainty), **CVaR** (risk-sensitive RL), **RAA** (risk-aware planning), and **HEU** (handcrafted policy that switches to GPS near obstacles). All methods share the same environment observations and cost structure, providing a fair evaluation of task-specific uncertainty handling.

Performance Comparison. Table 1 summarizes Task Completion Rate (TCR) and reward (R) for four representative tasks: *Goal Reaching (GR)*, *Avoid (AV)*, *Perimeter (PT)*, and *Multi-Goal (MG)*. G-SAC achieves the highest TCR and reward in all tasks, demonstrating safer navigation, fewer collisions, and more cost-effective use of precise GPS. Figure 1 (right) illustrates how G-SAC requests exact localization only in areas where the Task-Specific Uncertainty Map (TSUM) dictates tighter uncertainty requirements.

	GR		**AV**		**PT**		**MG**	
Method	**TCR**	**R**	**TCR**	**R**	**TCR**	**R**	**TCR**	**R**
SAC	68.9	144.1	71.3	177.8	44.3	84.4	31.3	124.4
SAC-P	84.3	241.8	83.2	199.2	51.6	132.8	42.9	135.2
B-SAC	74.2	189.2	79.6	277.6	56.3	170.4	37.7	122.4
CVaR	66.8	134.6	78.4	220.4	41.6	32.8	30.9	129.2
RAA	35.3	26.9	51.3	107.8	39.8	111.2	19.5	100.4
HEU	71.3	176.1	62.4	174.4	49.6	146.8	42.1	155.2
G-SAC	**92.0**	**410.4**	**88.7**	**482.2**	**83.7**	**594.6**	**81.7**	**511.4**

Table 1: Task Completion Rate (TCR) and average reward (R) for different tasks. G-SAC outperforms all baselines.

ACKNOWLEDGMENTS

This work is supported in part by NSF grants IIS-2024733 and IIS-2331908, the ONR grants N00014-23-1-2789, N00014-23-1-2651, and N00014-23-1-2505, the DHS grant 23STSLA00016-01-00, the DoD grant 78170-RT-REP, and the Florida Department of Environmental Protection grant INV31.

REFERENCES

[1] Mohammad Alali and Mahdi Imani. 2024. Bayesian reinforcement learning for navigation planning in unknown environments. *Frontiers in Artificial Intelligence* (2024).

[2] Matthew Budd, Paul Duckworth, Nick Hawes, and Bruno Lacerda. 2023. Bayesian reinforcement learning for single-episode missions in partially unknown environments. In *Conference on Robot Learning*.

[3] Matthew Chang, Theophile Gervet, Mukul Khanna, Sriram Yenamandra, Dhruv Shah, So Yeon Min, Kavit Shah, Chris Paxton, Saurabh Gupta, Dhruv Batra, et al. 2023. GOAT: Go to any thing. *arXiv preprint arXiv:2311.06430* (2023).

[4] Tao Chen, Saurabh Gupta, and Abhinav Gupta. 2019. Learning exploration policies for navigation. *arXiv preprint arXiv:1903.01959* (2019).

[5] Gerald Cook and Feitian Zhang. 2020. *Mobile robots: Navigation, control and sensing, surface robots and AUVs.* John Wiley & Sons.

[6] Aidan Curtis, George Matheos, Nishad Gothoskar, Vikash Mansinghka, Joshua Tenenbaum, Tomás Lozano-Pérez, and Leslie Pack Kaelbling. 2024. Partially Observable Task and Motion Planning with Uncertainty and Risk Awareness. *arXiv preprint arXiv:2403.10454* (2024).

[7] Dieter Fox, Sebastian Thrun, Wolfram Burgard, and Frank Dellaert. 2001. Particle filters for mobile robot localization. In *Sequential Monte Carlo methods in practice*.

[8] Alejandro Gonzalez-Garcia and Herman Castañeda. 2021. Guidance and control based on adaptive sliding mode strategy for a USV subject to uncertainties. *IEEE Journal of Oceanic Engineering* (2021).

[9] Faiza Gul, Wan Rahiman, and Syed Sahal Nazli Alhady. 2019. A comprehensive study for robot navigation techniques. *Cogent Engineering* (2019).

[10] Siyu Guo, Xiuguo Zhang, Yiquan Du, Yisong Zheng, and Zhiying Cao. 2021. Path planning of coastal ships based on optimized DQN reward function. *Journal of Marine Science and Engineering* (2021).

[11] Fredrik Gustafsson, Fredrik Gunnarsson, Niclas Bergman, Urban Forssell, Jonas Jansson, Rickard Karlsson, and P-J Nordlund. 2002. Particle filters for positioning, navigation, and tracking. *IEEE Transactions on signal processing* (2002).

[12] Tuomas Haarnoja, Aurick Zhou, Pieter Abbeel, and Sergey Levine. 2018. Soft actor-critic: Off-policy maximum entropy deep reinforcement learning with a stochastic actor. In *International Conference on Machine Learning*.

[13] Astghik Hakobyan, Gyeong Chan Kim, and Insoon Yang. 2019. Risk-aware motion planning and control using CVaR-constrained optimization. *IEEE Robotics and Automation letters* (2019).

[14] Jonathan Ko and Dieter Fox. 2009. GP-BayesFilters: Bayesian filtering using Gaussian process prediction and observation models. *Autonomous Robots* (2009).

[15] Aviral Kumar, Justin Fu, Matthew Soh, George Tucker, and Sergey Levine. 2019. Stabilizing off-policy Q-Learning via bootstrapping error reduction. *Advances in Neural Information Processing Systems* (2019).

[16] Guofa Li, Yifan Yang, Shen Li, Xingda Qu, Nengchao Lyu, and Shengbo Eben Li. 2022. Decision making of autonomous vehicles in lane change scenarios: Deep reinforcement learning approaches with risk awareness. *Transportation research part C: emerging technologies* (2022).

[17] Yueming Li, Mingquan Ma, Jian Cao, Guobin Luo, Depeng Wang, and Weiqiang Chen. 2024. A Method for Multi-AUV Cooperative Area Search in Unknown Environment Based on Reinforcement Learning. *Journal of Marine Science and Engineering* (2024).

[18] Yinhan Liu. 2019. Roberta: A robustly optimized BERT pretraining approach. *arXiv preprint arXiv:1907.11692* (2019).

[19] Steven Morad, Ryan Kortvelesy, Matteo Bettini, Stephan Liwicki, and Amanda Prorok. 2023. POPGym: Benchmarking partially observable reinforcement learning. *arXiv preprint arXiv:2303.01859* (2023).

[20] Masahiro Ono, Thoams J Fuchs, Amanda Steffy, Mark Maimone, and Jeng Yen. 2015. Risk-aware planetary rover operation: Autonomous terrain classification and path planning. In *IEEE Aerospace Conference*.

[21] Abhishek Padalkar, Acorn Pooley, Ajinkya Jain, Alex Bewley, Alex Herzog, Alex Irpan, Alexander Khazatsky, Anant Rai, Anikait Singh, Anthony Brohan, et al. 2023. Open X-Embodiment: Robotic learning datasets and RT-X models. *arXiv preprint arXiv:2310.08864* (2023).

[22] BK Patle, Anish Pandey, DRK Parhi, AJDT Jagadeesh, et al. 2019. A review: On path planning strategies for navigation of mobile robot. *Defence Technology* (2019).

[23] Gokul Puthumanaillam, Xiangyu Liu, Negar Mehr, and Melkior Ornik. 2024. Weathering ongoing uncertainty: Learning and planning in a time-varying partially observable environment. In *IEEE International Conference on Robotics and Automation*.

[24] Gokul Puthumanaillam, Jae Hyuk Song, Nurzhan Yesmagambet, Shinkyu Park, and Melkior Ornik. 2024. TAB-Fields: A Maximum Entropy Framework for Mission-Aware Adversarial Planning. *arXiv preprint arXiv:2412.02570* (2024).

[25] Gokul Puthumanaillam, Manav Vora, and Melkior Ornik. 2024. ComTraQ-MPC: Meta-Trained DQN-MPC Integration for Trajectory Tracking with Limited Active Localization Updates. *arXiv preprint arXiv:2403.01564* (2024).

[26] Krishan Rana, Ben Talbot, Vibhavari Dasagi, Michael Milford, and Niko Sünderhauf. 2020. Residual reactive navigation: Combining classical and learned navigation strategies for deployment in unknown environments. In *2020 IEEE International Conference on Robotics and Automation*.

[27] Cesar A Rojas, Paulo Padrão, Jose Fuentes, Gregory M Reis, Arif R Albayrak, Batuhan Osmanoglu, and Leonardo Bobadilla. 2024. Combining multi-satellite remote and in-situ sensing for unmanned underwater vehicle state estimation. *Ocean Engineering* (2024).

[28] Dhruv Shah, Błażej Osiński, Sergey Levine, et al. 2023. LM-Nav: Robotic navigation with large pre-trained models of language, vision, and action. In *Conference on Robot Learning*.

[29] Siddharth Singi, Zhanpeng He, Alvin Pan, Sandip Patel, Gunnar A Sigurdsson, Robinson Piramuthu, Shuran Song, and Matei Ciocarlie. 2024. Decision making for human-in-the-loop robotic agents via uncertainty-aware reinforcement learning. In *2024 IEEE International Conference on Robotics and Automation*.

[30] Matthijs TJ Spaan, Tiago S Veiga, and Pedro U Lima. 2015. Decision-theoretic planning under uncertainty with information rewards for active cooperative perception. *Autonomous Agents and Multi-Agent Systems* (2015).

[31] Krishnan Srinivasan, Benjamin Eysenbach, Sehoon Ha, Jie Tan, and Chelsea Finn. 2020. Learning to be safe: Deep RL with a safety critic. *arXiv preprint arXiv:2010.14603* (2020).

[32] Cyrill Stachniss, Wolfram Burgard, et al. 2014. Particle filters for robot navigation. *Foundations and Trends in Robotics* (2014).

[33] Annalisa T Taylor, Thomas A Berrueta, and Todd D Murphey. 2021. Active learning in robotics: A review of control principles. *Mechatronics* (2021).

[34] Ning Wang, Yabiao Wang, Yuming Zhao, Yong Wang, and Zhigang Li. 2022. Sim-to-real: Mapless navigation for USVs using deep reinforcement learning. *Journal of Marine Science and Engineering* (2022).

[35] Zixiang Wang, Hao Yan, Zhuoyue Wang, Zhengjia Xu, Zhizhong Wu, and Yining Wang. 2024. Research on autonomous robots navigation based on reinforcement learning. In *International Conference on Robotics, Artificial Intelligence and Intelligent Control*. IEEE.

[36] Zichao Yang, Diyi Yang, Chris Dyer, Xiaodong He, Alex Smola, and Eduard Hovy. 2016. Hierarchical attention networks for document classification. In *Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*.

[37] Denis Yarats, Amy Zhang, Ilya Kostrikov, Brandon Amos, Joelle Pineau, and Rob Fergus. 2021. Improving sample efficiency in model-free reinforcement learning from images. In *AAAI Conference on Artificial Intelligence*.

[38] Yu Zhang, Sarath Sreedharan, Anagha Kulkarni, Tathagata Chakraborti, Hankz Hankui Zhuo, and Subbarao Kambhampati. 2016. Plan explicability for robot task planning. In *RSS Workshop on Planning for Human-Robot Interaction: Shared Autonomy and Collaborative Robotics*.

[39] Kai Zhu and Tao Zhang. 2021. Deep reinforcement learning based mobile robot navigation: A review. *Tsinghua Science and Technology* (2021).

[40] Yuanyang Zhu, Zhi Wang, Chunlin Chen, and Daoyi Dong. 2021. Rule-based reinforcement learning for efficient robot navigation with space reduction. *IEEE/ASME Transactions on Mechatronics* (2021).