Surrogate Assisted Generation of Human-Robot Interaction Scenarios

Anonymous Author(s)
Affiliation
Address
email

Abstract: As human-robot interaction (HRI) systems advance, so does the difficulty of evaluating and understanding the strengths and limitations of these systems in different environments and with different users. To this end, previous methods have algorithmically generated diverse scenarios that reveal system failures in a shared control teleoperation task. However, these methods require directly evaluating generated scenarios by simulating robot policies and human actions. The computational cost of these evaluations limits their applicability in more complex domains. Thus, we propose augmenting scenario generation systems with surrogate models that predict both human and robot behaviors. In the shared control teleoperation domain and a more complex shared workspace collaboration task, we show that surrogate assisted scenario generation efficiently synthesizes diverse datasets of challenging scenarios. We demonstrate that these failures are reproducible in real-world interactions.

Keywords: Scenario Generation, Human-Robot Interaction, Quality Diversity

1 Introduction

As the complexity of robotic systems that interact with people increases, it becomes impossible for designers and end-users to anticipate how a robot will act in different environments and with different users. For instance, consider a robotic arm collaborating with a user on a package labeling task, where a user attaches a label while the robot presses a stamp with the goal of completing the task as fast as possible (Fig. 1). In this task, the arm infers the user's intended goal object and moves simultaneously towards a different object to avoid collision. The robot's motion depends on which object the user selects to label, how the user moves towards that object, and how all objects are arranged in the environment. Thus, evaluating the system requires testing it with a diverse range of user behaviors and object arrangements.

While user studies are essential for understanding how users will interact with a robot, they are limited in the number of environments and user behaviors they can cover. This makes algorithmic scenario generation a compelling method that can complement user studies by finding failures and elucidating a holistic view of the strengths and limitations of a robotic system's behavior.

Our goal is not to find a single maximally adversarial scenario but to generate a diverse dataset of challenging scenarios. In our example package labeling task, delays in task execution may occur due to incorrect inference of the user's goal, how the goal objects are arranged in the environment, and how the user and the robot move towards these objects.

Previous work [1, 2] has formulated scenario generation as a quality diversity (QD) problem and demonstrated the effectiveness of QD algorithms in generating diverse collections of scenarios in a shared control teleoperation domain. In that domain, a user teleoperates a robotic arm with a joystick interface, while the robot observes the joystick inputs to infer the user's goal and assist the user in reaching their goal.

Submitted to the 7th Conference on Robot Learning (CoRL 2023). Do not distribute.

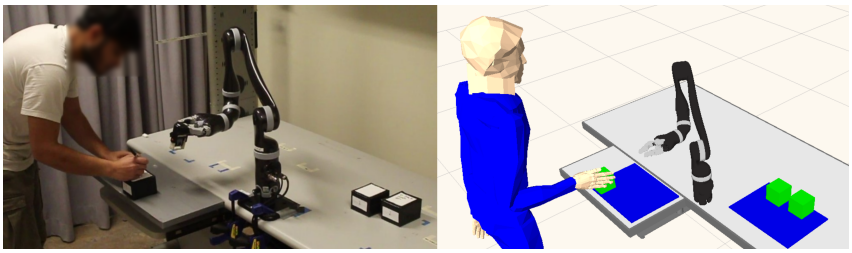


Figure 1: Example scenario in a collaborative package labeling task found by our proposed surrogate assisted scenario generation framework. The presence of the two objects behind the robot results in its expected cost-minimizing policy to move towards the object in the front, resulting in a conflict with the user who is reaching the object at the same time.

While prior work could generate diverse datasets of failure scenarios in one-shot shared teleoperation interactions, these only last a few seconds in contrast to collaborative, sequential tasks that last much longer. For instance, completing the package labeling task (Fig. 1) can take several minutes. The long duration of such human-robot interaction scenarios makes their evaluation expensive. This limits the applicability of QD algorithms, which require a large number of evaluations [3].

Our key insight is that *we can train deep neural networks as surrogate models to predict human-robot interaction outcomes and integrate them into the scenario generation process*. In addition to making scenario evaluations inexpensive, deep neural networks are end-to-end differentiable, which allows us to integrate state-of-the-art differentiable quality diversity (DQD) algorithms [4] to efficiently discover scenarios that the surrogate model predicts are challenging with diverse behavior.

We make the following contributions: (1) We introduce using deep neural networks as surrogate models to predict human-robot interaction outcomes, such as time to task completion, maximum robot path length, or total waiting time; (2) We integrate surrogate models with differentiable quality diversity (DQD) algorithms that leverage gradient information backpropagated through the surrogate model. (3) We show, in the shared control teleoperation domain of previous work [1] and in a shared workspace collaboration domain, that surrogate assisted scenario generation results in significant benefits in terms of sample efficiency. It also achieves a significant reduction in computational time in the collaboration domain, where evaluations are particularly expensive.

2 Problem Statement

We model the problem of generating a diverse and challenging dataset of human-robot interaction scenarios as a quality diversity (QD) problem and adopt the QD definition from prior work [4].

We assume a scenario parameterized by $\boldsymbol{\theta} \in \mathbb{R}^n$. The scenario parameters could be object positions and types in the environment, human model parameters, or latent inputs to a generative model of environments, which is converted to a scenario via a function $G(\boldsymbol{\theta})$. The objective function $f : \mathbb{R}^n \to \mathbb{R}$ assesses the quality of a scenario $\boldsymbol{\theta}$. Because we wish to find challenging scenarios, higher quality implies worse team performance, e.g., longer task completion time.

We further assume a set of user-defined measure functions, $m_i : \mathbb{R}^n \to \mathbb{R}$, or as a vector function $\boldsymbol{m} : \mathbb{R}^n \to \mathbb{R}^k$, that quantify aspects of the scenario that we wish to diversify, e.g., distance between objects, noise in user inputs, or human and robot path length. The range of $\boldsymbol{m}$ forms a measure space $S = \boldsymbol{m}(\mathbb{R}^n)$, which we assume is tessellated into M cells, forming an *archive*.

The QD objective is to maximize the QD-score [5]:

$$\max_{\boldsymbol{\theta}_i} \sum_{i=1}^{M} f(\boldsymbol{\theta}_i) \tag{1}$$

Here $\boldsymbol{\theta}_i$ refers to the scenario with the highest quality in cell i of the archive. If there are no scenarios in a cell, $f(\boldsymbol{\theta}_i)$ is assumed to be zero.

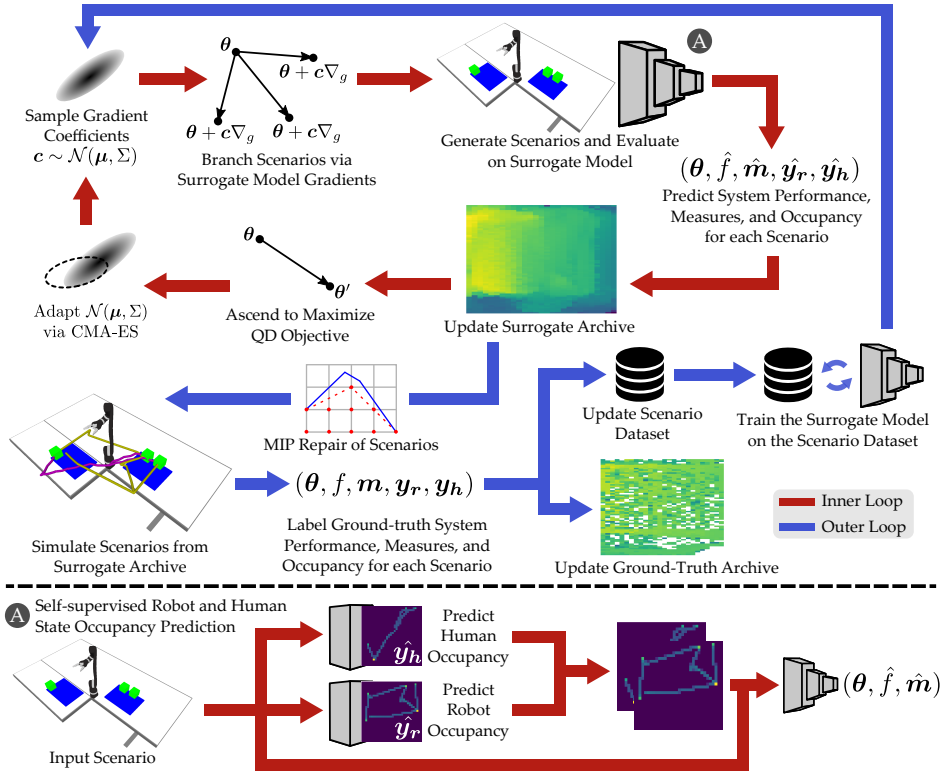


Figure 2: An overview of our proposed differentiable surrogate assisted scenario generation (DSAS) algorithm for HRI tasks. The algorithm runs in two stages: an inner loop to exploit a surrogate model of the human and the robot behavior (**red arrows**) and an outer loop to evaluate candidate scenarios and add them to a dataset (**blue arrows**). See App. A for the complete pseudocode.

The differentiable quality diversity (DQD) problem formulation is a special case of QD where the objective function f and measure functions $\boldsymbol{m}$ are first-order differentiable.

3 Background

Scenario Generation. Algorithmic scenario generation has many applications, which include designing video game levels [6, 7, 8, 9, 10, 11, 12] and testing autonomous vehicles [13, 14, 15, 16, 17, 18, 19], motion planning algorithms [20], and reinforcement learning agents [21, 22, 23, 24]. It has also been applied to create curricula for robot learning [25, 26, 27, 28] and to co-evolve agents and environments for agent generalizability [29, 30, 31, 32, 33, 34, 35, 36, 37, 38]. Most relevant to our work is prior work [1, 2] in human-robot interaction that applied the MAP-Elites [39] and CMA-ME [40] quality diversity (QD) algorithms to find robot failures in a shared control teleoperation domain. In this work, we significantly improve sample and wall-clock efficiency by combining the state-of-the-art QD algorithms CMA-MAE [41] and CMA-MAEGA [41] with surrogate models that predict human-robot interaction outcomes. We test our approach both in the shared control teleoperation domain of previous work [1] and in a shared workspace human-robot collaboration domain, originally proposed in previous work on human-robot teaming [42].

QD Algorithms. QD algorithms, such as MAP-Elites [39] and CMA-MAE [40, 41], solve the QD problem defined in Sec. 2. They have been used to generate diverse locomotion strategies [39, 43], video game levels [44], nano-materials [45], and building layouts [46]. Certain prior QD algorithms [3, 47, 48, 21] have leveraged surrogate models based on Gaussian processes or deep neural networks to guide the search. Most relevant to our work is the Deep Surrogate Assisted Generation of Environments (DSAGE) [21] algorithm, which exploits a surrogate model with quality diversity

algorithms to generate environments. However, DSAGE has only been applied in *single-agent grid-world* game domains. In contrast, this work addresses the much more complex task of human-robot interaction scenarios, which requires the advances in the surrogate model, QD search, and scenario generation described in Sec. 4.

4 Surrogate Assisted Scenario Generation

Our method for algorithmically generating diverse collections of HRI scenarios builds upon recent work in generating single-agent grid-world environments with surrogate models [21]. We briefly describe the advances necessary to scale these techniques from single-agent grid-world game domains to the much more complex HRI domains.

Surrogate Models for Human-Robot Interaction. We scale the surrogate model to predict outcomes of HRI scenarios that include both robot and human behavior and environment parameters. First, we allow both the environment and the human model parameters as inputs to the surrogate model. Second, we discretize the shared workspace and predict two occupancy grids, one for the human and one for the robot. We then stack both predictions as inputs to a convolutional neural network, which predicts the objective and measure functions. App. B provides precise details of our surrogate model.

Scenario Repair via Mixed Integer Programming. In contrast to previous work [1] that considered a single workspace, scenarios in our work have disjoint workspaces, with each workspace imposing constraints on object arrangement, such as boundary and collision avoidance constraints. We thus adopt a generate-then-repair approach [49, 50], where we generate unconstrained scenarios and pass them through a mixed integer program (MIP). We propose a MIP formulation that solves for the minimum cost edit of object locations – the sum of object displacement distances – that satisfies the constraints for a valid scenario. See App. C for the complete MIP formulation.

Objective Regularization. The generate-then-repair strategy repairs invalid scenarios by moving objects placed outside the workspace boundaries to the edge of the workspaces. However, this does not incentivize the QD algorithm to search the workspace interiors, which can result in the search diverging away from the workspace areas. To guide the search towards the workspace interiors, we discount the objective function f of the QD formulation by the cost of the MIP repair. An ablation study in App. G.3 shows the effects of objective regularization on performance.

DQD with Surrogate Models. The DSAGE algorithm exploits the surrogate model with *derivative-free* QD algorithms. A key observation is that the surrogate model is an end-to-end differentiable neural network. We can take advantage of this by exploiting the surrogate model with differentiable quality diversity (DQD) algorithms [4], which leverage the gradients of the objective and measure functions to accelerate QD optimization.

Algorithm. Fig. 2 provides an overview of the complete algorithm, which consists of an outer loop (blue arrows) and an inner loop (red arrows). In the inner loop, a QD algorithm searches for scenarios that are challenging and diverse according to the surrogate model predictions. We repair the generated scenarios to ensure validity and evaluate each repaired scenario to obtain ground truth objective and measure values. Based on these values, we add each scenario to a ground-truth archive and a training dataset for the surrogate model. The surrogate model trains on this dataset, correcting prediction errors exploited by the QD algorithm. After accumulating enough diverse data, the surrogate model starts making accurate predictions, and the inner loop produces truly diverse and challenging scenarios. After multiple outer loop iterations, the ground-truth archive accumulates diverse and challenging scenarios, testing the HRI system's strengths and limitations.

The key idea behind the proposed algorithm is that exploiting the surrogate model with QD algorithms produces diverse – with respect to the surrogate model predictions – scenarios. Labeling these scenarios by evaluating them in a simulator and using them to retrain the surrogate model in turn improves its predictions in subsequent iterations. Thus, the surrogate model self-improves over time by training on the diverse data generated according to its predictions.

The proposed improvements result in two versions of our algorithm. The derivative-free version, Surrogate Assisted Scenario Generation (SAS), employs a derivative-free QD algorithm, CMA-MAE, in the inner loop. The differentiable version, Differentiable Surrogate Assisted Scenario Generation (DSAS, Fig. 2), employs a DQD algorithm, CMA-MAEGA, in the inner loop. Additional algorithm details and the pseudocode are given in App. A.

5 Domains

We consider two HRI domains from prior work: shared control teleoperation [51] and shared workspace collaboration [42] with a 6-DoF Gen2 Kinova JACO arm.

Shared Control Teleoperation. In shared control teleoperation, a user provides low-dimensional joystick inputs towards a goal. To aid the user, the robot attempts to infer the human goal and move towards it autonomously. The robot solves a POMDP with the user's goal as a latent variable and updates its belief based on the human input, assuming a noisily-optimal user [51]. With hindsight optimization and first-order approximation, this results in the robot's actions being a weighted average of the optimal path towards each goal, where the weights are proportional to the respective goal probabilities. In App. G.1, we test a second robot policy that blends the user's and the robot's actions based on the robot's confidence in the user's goal [52]. Scenario parameters include both the environment, i.e., the coordinates of two goal objects, and the human actions, i.e., a set of human trajectory waypoints. To search for failures, we set the objective to be the time taken to reach the correct goal. We diversify scenarios with respect to the noise in human inputs (variation from the optimal path) and the scene clutter (distance between goals).

Shared Workspace Collaboration. In shared workspace collaboration, the human and the robot simultaneously execute a sequential task in a shared workspace with disjoint regions [42, 53]. As in the teleoperation task, the robot models the human goal as a latent variable in a POMDP and uses hindsight optimization and first-order value function approximations to act in real time. However, in this task, the human acts independently, and the robot tracks the position of the human hand as its observation. Furthermore, the robot attempts to avoid the goal intended by the human, selecting the nearest goal from a feasible set of goals different than the user's, i.e., it maps a human candidate goal to a different goal-to-go. The robot's action is a weighted average of the optimal path towards each goal-to-go, with weights proportional to the probability of the corresponding human goal.

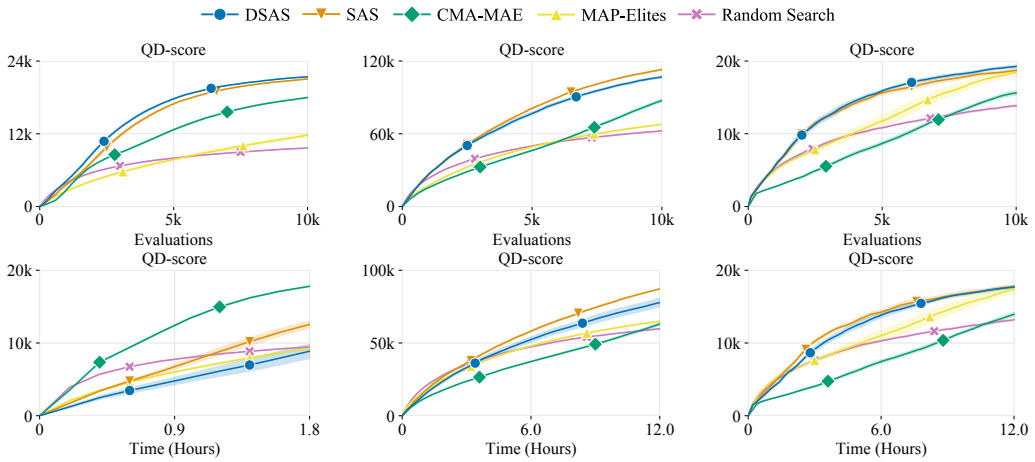
We parameterize the scenario with three goal coordinates and set the objective to be the task completion time. We select measures based on factors that we expect to affect the team performance: object arrangement, the accuracy of inference of user's goal, and the distance required to reach the goals. We choose two sets of measures: (1) The minimum distance between goal objects and the maximum probability assigned by the robot to the wrong goal. (2) The robot path length and the total time for which the human and robot have to wait when reaching the same goal.

Unlike the teleoperation domain where we search directly over human actions (similar to prior work [1, 2]), in this domain we adopt a more general model of the human as solving a softmax MDP with the maximum entropy formulation [54]. In App. G.2, we include an additional setting in the collaboration domain, where we search over both environments and human model parameters related to speed and rationality. Additional details about the domains, the robot policy, the human policy, and the implementation are provided in App. D, App. E.1, App. E.2, and App. F respectively.

6 Experiments

Independent Variables. Our two independent variables are the domain and the algorithm.

Our three domains are: (a) shared control teleoperation with distance between the goals and human variation as measures; (b) shared workspace collaboration with minimum distance between the goals and maximum wrong goal probability as measures (collaboration I); (c) shared workspace collaboration with robot path length and total wait time as measures (collaboration II).



(a) Shared control teleoperation (b) Collaboration I (c) Collaboration II

Figure 3: QD-score attained in the three domains as a function of the number of evaluations (top) and the wall-clock time (bottom). Plots show the mean and standard error of the mean.

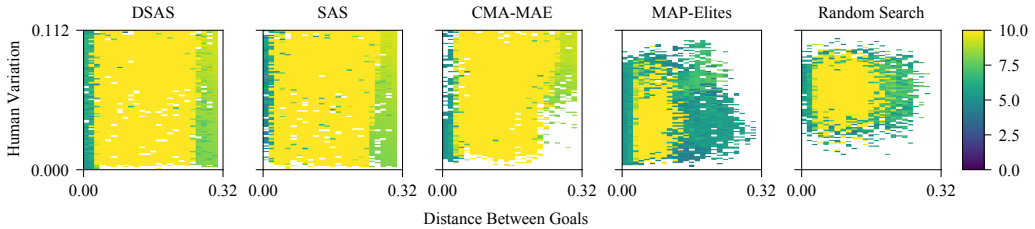


Figure 4: Comparison of the final archive heatmaps in the shared control teleoperation domain.

In each domain, we compare five different algorithms: (a) **Random Search**, where we uniformly sample scenarios from the valid regions. (b) **MAP-Elites** [39], as adapted for scenario generation in previous work [1], with the additional objective regularization described in Sec. 4. (c) **CMA-MAE** [41] with objective regularization. (d) **SAS**: The proposed derivative-free version of our surrogate assisted scenario generation algorithm. We apply CMA-MAE as the derivative-free QD algorithm in the inner loop. (e) **DSAS**: The proposed differentiable surrogate assisted scenario generation with the DQD algorithm CMA-MAEGA in the inner loop.

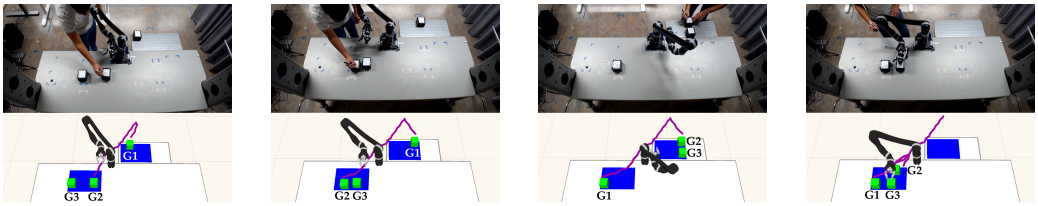
Dependent Variable: Following previous work [1], we set QD-score [5] (Eq. 1 in Sec. 2) as the dependent variable that summarizes the quality and diversity of solutions. We compute the QD-score at the end of 10,000 evaluations, averaged over 10 trials of random search, MAP-Elites, CMA-MAE, and – because of GPU usage constraints – 5 trials of SAS and DSAS.

Hypotheses:

H1. *We hypothesize that the surrogate assisted QD algorithms SAS and DSAS will outperform CMA-MAE, MAP-Elites, and random search.* We base this hypothesis on previous work, which has shown the benefit of integrating QD with surrogate models in single-agent domains [48, 21].

H2. *We hypothesize that DSAS will outperform SAS.* We base this hypothesis on previous work, which has shown that DQD algorithms perform significantly better than their derivative-free counterparts [4] when the objective and measure gradients are available.

Analysis. A two-way ANOVA test showed a significant interaction effect ($F(8.0, 105.0) = 305.79, p < 0.001$). Simple main effects analysis on each domain showed a significant effect of the algorithm on the QD-score ($p < 0.001$).



(a) Scenario 1 (b) Scenario 2 (c) Scenario 3 (d) Scenario 4

Figure 5: Example scenarios recreated with a real robot. The purple line shows the simulated human path. Videos of the simulated and recreated scenarios are included in the supplemental material.

Pairwise t-tests with Bonferroni corrections showed that SAS and DSAS performed significantly better than CMA-MAE, MAP-Elites, and random search in the shared control teleoperation and collaboration I domains ($p < 0.001$). However, in the collaboration II domain, they outperformed CMA-MAE ($p < 0.001$) and random search ($p < 0.001$), while there was no significant difference with MAP-Elites. We attribute this to the large variance of the MAP-Elites runs and to the fact that MAP-Elites can easily populate the archive for the given measure of the robot's path length by making small isotropic perturbations in the object positions (see App. G for details). Fig. 3 shows the QD-score as a function of the number of evaluations. We see that both SAS and DSAS achieve a high QD-score early in the search, indicating high sample efficiency.

The comparison between SAS and DSAS showed mixed results, with SAS performing significantly better in the collaboration I domain ($p < 0.001$), DSAS performing significantly better in the shared control teleoperation domain ($p < 0.001$), and no significance in the collaboration II domain ($p = 0.07$). We observe that the shared control teleoperation domain has a 9-dimensional search space, while the shared workspace collaboration task has a 6-dimensional search space. Previous work [4, 41] has shown DQD algorithms improving efficiency in very high-dimensional search spaces by reducing the search from a high-dimensional solution space to a low-dimensional objective-measure space. We conjecture that this explains the significant improvement in the shared control teleoperation domain and we will investigate higher-dimensional domains in future work.

Furthermore, in the shared workspace collaboration domains, where scenario evaluations last a couple of minutes because of the larger task complexity, surrogate assistance showed wall-clock time efficiency (Fig. 3), unlike prior work [21] that only showed sample efficiency improvements.

Fig. 4 shows heatmaps of the final archives in the shared control teleoperation domain. The heatmaps for MAP-Elites and random search match the results from prior work [1]. SAS and DSAS nearly fill in the archive, including scenarios in the lower right corner of the archive. This is the hardest region of the archive to find failures because it includes scenarios with a nearly optimal human and a large distance between the goal objects, which would typically enable correct inference of the human goal. App. G includes heatmaps in other domains and tabulates the QD-scores.

We provide additional experiments with two different settings in App. G.1 and App. G.2. Additionally, we show the effect of objective regularization in App. G.3 and examples of generated scenarios with high (as opposed to low) team performance in App. I.

Real World Demo. We wish to demonstrate that the generated failure scenarios are reproducible in the real world and are not just simulation artifacts. Thus, we recreate four example scenarios from the generated archives with a 6-DoF Gen2 Kinova JACO arm and by having a user reproduce the motions of the simulated human. We track the human hand position with a Kinect v1 sensor and the OpenNI package [55] for skeleton tracking. We discuss two of the scenarios below and the other two in App. H. We include videos of all four scenarios in the supplementary material.

Incorrect robot motion because of delayed human goal inference (Fig. 5a): We select this scenario from the archive generated by DSAS in the collaboration II domain. The scenario time was 81s and the robot path length was 4.9m. In this scenario, after the human finishes working on goal G1 and the robot on G3, the robot is closer to G2 than its other remaining goal, G1. Based on the feasible goal

set formulation (Sec. 5), G2 becomes the goal-to-go for human candidate goals G1 and G3. Given that the combined probability of the human going to either G1 or G3 is higher than the probability of the human going to G2, the robot moves towards G2. However, once the robot realizes that the human is actually moving to G2 as well, the robot has to move all the way back to goal G1, resulting in a significant delay.

Long robot motion with correct human goal inference (Fig. 5c): We additionally wish to find scenarios that result in poor team performance that is not due to incorrect inference. We select a scenario from a SAS archive in the collaboration I domain that had a completion time of 74s and a low maximum wrong goal probability of 0.3. The poor performance here is caused by the interaction between the robot's policy and the object placement. As the robot moves between the two workspaces following a straight line path, the robot reaches a configuration close to self-collision or to joint limits, which prompts the system to perform a re-planning motion back to its start configuration. This happened twice during task execution, resulting in a significant delay.

7 Discussion

Limitations. Our approach scales surrogate assisted scenario generation from single-agent grid-world domains to complex human-robot interaction domains with continuous actions, environment dynamics, and object locations. However, our evaluation domains consist of objects of the same type and simple human models. We note that SAS and DSAS are general algorithms and we are excited about integrating them with more complex models of environments [50] and human actions [56, 57], as well as leveraging high-fidelity human model simulators [58, 59] to improve realism. Furthermore, our system does not explain the reason behind the observed robot behavior in the generated scenarios, and future work will explore integrating scenario generation with methods for failure explanation [60, 61]. Finally, while we focus on a single human interacting with a single robot, we believe that our workspace occupancy-based approach for surrogate model predictions can be extended to multi-human-robot team settings.

Implications. We presented the SAS and DSAS scenario generation algorithms that accelerate QD scenario generation via surrogate models. Results in a shared control teleoperation domain of previous work [1] and in a shared workspace collaboration domain show significant improvements in search efficiency when generating diverse datasets of challenging scenarios.

For the first time in surrogate assisted scenario generation methods, we see improvements not only in sample efficiency but also in wall-clock time in the shared workspace collaboration domain, where evaluations last a couple of minutes. On the other hand, the additional computation in the inner loop of the surrogate assisted algorithms resulted in more time required to match and exceed the performance of the baselines in the shared control teleoperation domain, where scenario evaluations last only a few seconds. Thus, for running-time performance, we recommend surrogate assisted methods in domains with expensive evaluations, in which the additional computation in the inner loop is offset by the improvement in sample efficiency.

We additionally highlight an unexpected benefit of our system during development. When we tested the shared workspace collaboration domain, SAS and DSAS discovered failure scenarios that exploited bugs in our implementation which were subsequently fixed. For instance, some goal locations were reachable by the robot arm in the real world but unreachable in simulation because of small errors in the robot's URDF file, which prompted us to correct it.

Overall, we envision the proposed algorithms as a valuable tool to accelerate the development and testing of HRI systems before user studies and deployment. We consider this an important step towards circumventing costly failures and reducing the risk of human injuries, which is a critical milestone for widespread acceptance and use of HRI systems.

References

[1] M. C. Fontaine and S. Nikolaidis. A quality diversity approach to automatically generating human-robot interaction scenarios in shared autonomy. In *Robotics: Science and Systems*, 2021. doi:10.15607/RSS.2021.XVII.036.

[2] M. C. Fontaine and S. Nikolaidis. Evaluating human–robot interaction algorithms in shared autonomy via quality diversity scenario generation. *ACM Transactions on Human-Robot Interaction (THRI)*, 11(3):1–30, 2022.

[3] A. Gaier, A. Asteroth, and J.-B. Mouret. Data-efficient design exploration through surrogate-assisted illumination. *Evolutionary Computation*, 2018.

[4] M. C. Fontaine and S. Nikolaidis. Differentiable quality diversity. In *Advances in Neural Information Processing Systems*, 2021. URL https://proceedings.neurips.cc/paper/2021/file/532923f11ac97d3e7cb0130315b067dc-Paper.pdf.

[5] J. K. Pugh, L. B. Soros, P. A. Szerlip, and K. O. Stanley. Confronting the challenge of quality diversity. In *Proceedings of the Annual Conference on Genetic and Evolutionary Computation*, 2015.

[6] D. Gravina, A. Khalifa, A. Liapis, J. Togelius, and G. N. Yannakakis. Procedural content generation through quality diversity. In *Proceedings of the IEEE Conference on Games (CoG)*, 2019.

[7] M. C. Fontaine, R. Liu, A. Khalifa, J. Modi, J. Togelius, A. K. Hoover, and S. Nikolaidis. Illuminating mario scenes in the latent space of a generative adversarial network. In *Proceedings of the Thirty-Fifth AAAI Conference on Artificial Intelligence*, 2021. URL https://ojs.aaai.org/index.php/AAAI/article/view/16740.

[8] S. Earle, J. Snider, M. C. Fontaine, S. Nikolaidis, and J. Togelius. Illuminating diverse neural cellular automata for level generation. *CoRR*, abs/2109.05489, 2021. URL https://arxiv.org/abs/2109.05489.

[9] A. Khalifa, S. Lee, A. Nealen, and J. Togelius. Talakat: Bullet hell generation through constrained map-elites. In *Proceedings of the Genetic and Evolutionary Computation Conference*, 2018.

[10] K. Steckel and J. Schrum. Illuminating the space of beatable lode runner levels produced by various generative adversarial networks. In *Proceedings of the Genetic and Evolutionary Computation Conference*, 2021. doi:10.1145/3449726.3459440.

[11] J. Schrum, V. Volz, and S. Risi. CPPN2GAN: combining compositional pattern producing networks and gans for large-scale pattern generation. In *Proceedings of the Genetic and Evolutionary Computation Conference*, 2020. doi:10.1145/3377930.3389822.

[12] A. Sarkar and S. Cooper. Generating and blending game levels via quality-diversity in the latent space of a variational autoencoder. In *Proceedings of the 16th International Conference on the Foundations of Digital Games*, 2021. doi:10.1145/3472538.3472545.

[13] J. Arnold and R. Alexander. Testing autonomous robot control software using procedural content generation. In *Proceedings of the 32nd International Conference on Computer Safety, Reliability, and Security*, 2013. doi:10.1007/978-3-642-40793-2_4.

[14] G. E. Mullins, P. G. Stankiewicz, R. C. Hawthorne, and S. K. Gupta. Adaptive generation of challenging scenarios for testing and evaluation of autonomous vehicles. *Journal of Systems and Software*, 2018. doi:10.1016/j.jss.2017.10.031.

[15] Y. Abeysirigoonawardena, F. Shkurti, and G. Dudek. Generating adversarial driving scenarios in high-fidelity simulators. In *International Conference on Robotics and Automation, ICRA*, 2019. doi:10.1109/ICRA.2019.8793740. URL https://doi.org/10.1109/ICRA.2019.8793740.

[16] E. Rocklage, H. Kraft, A. Karatas, and J. Seewig. Automated scenario generation for regression testing of autonomous vehicles. In *Proceedings of the 20th IEEE International Conference on Intelligent Transportation Systems (ITSC)*, 2017. doi:10.1109/ITSC.2017.8317919.

[17] A. Gambi, M. Mueller, and G. Fraser. Automatically testing self-driving cars with search-based procedural content generation. In *Proceedings of the 28th ACM SIGSOFT International Symposium on Software Testing and Analysis*, 2019. doi:10.1145/3293882.3330566.

[18] D. Sadigh, S. S. Sastry, and S. A. Seshia. Verifying robustness of human-aware autonomous cars. *IFAC-PapersOnLine*, 2019.

[19] D. J. Fremont, T. Dreossi, S. Ghosh, X. Yue, A. L. Sangiovanni-Vincentelli, and S. A. Seshia. Scenic: a language for scenario specification and scene generation. In *Proceedings of the 40th ACM SIGPLAN Conference on Programming Language Design and Implementation*, 2019.

[20] Y. Zhou, S. Booth, N. Figueroa, and J. Shah. RoCUS: robot controller understanding via sampling. In *Proceedings of the Conference on Robot Learning*, 2021. URL https://proceedings.mlr.press/v164/zhou22a.html.

[21] V. Bhatt, B. Tjanaka, M. C. Fontaine, and S. Nikolaidis. Deep surrogate assisted generation of environments. *Conference on Neural Information Processing (NeurIPS)*, 2022.

[22] M. Deitke, E. VanderBilt, A. Herrasti, L. Weihs, J. Salvador, K. Ehsani, W. Han, E. Kolve, A. Farhadi, A. Kembhavi, et al. ProcTHOR: Large-scale embodied AI using procedural generation. *arXiv preprint arXiv:2206.06994*, 2022.

[23] K. Cobbe, C. Hesse, J. Hilton, and J. Schulman. Leveraging procedural generation to benchmark reinforcement learning. In *Proceedings of the International Conference on Machine Learning*, 2020.

[24] S. Risi and J. Togelius. Increasing generality in machine learning through procedural content generation. *Nature Machine Intelligence*, 2020.

[25] B. Mehta, M. Diaz, F. Golemo, C. J. Pal, and L. Paull. Active domain randomization. In *Conference on Robot Learning*, pages 1162–1176. PMLR, 2020.

[26] J. Lee, J. Hwangbo, L. Wellhausen, V. Koltun, and M. Hutter. Learning quadrupedal locomotion over challenging terrain. *Science robotics*, 2020.

[27] K. Fang, Y. Zhu, S. Savarese, and L. Fei-Fei. Discovering generalizable skills via automated generation of diverse tasks. *Robotics: Science and Systems*, 2021.

[28] K. Fang, T. Migimatsu, A. Mandlekar, L. Fei-Fei, and J. Bohg. Active task randomization: Learning visuomotor skills for sequential manipulation by proposing feasible and novel tasks. *arXiv preprint arXiv:2211.06134*, 2022.

[29] R. Wang, J. Lehman, J. Clune, and K. O. Stanley. Paired open-ended trailblazer (POET): endlessly generating increasingly complex and diverse learning environments and their solutions. *CoRR*, abs/1901.01753, 2019. URL http://arxiv.org/abs/1901.01753.

[30] R. Wang, J. Lehman, A. Rawal, J. Zhi, Y. Li, J. Clune, and K. O. Stanley. Enhanced POET: open-ended reinforcement learning through unbounded invention of learning challenges and their solutions. In *Proceedings of the 37th International Conference on Machine Learning, ICML*, 2020. URL http://proceedings.mlr.press/v119/wang20l.html.

[31] T. Gabor, A. Sedlmeier, M. Kiermeier, T. Phan, M. Henrich, M. Pichlmair, B. Kempter, C. Klein, H. Sauer, R. S. AG, et al. Scenario co-evolution for reinforcement learning on a grid world smart factory domain. In *Proceedings of the Genetic and Evolutionary Computation Conference*, 2019.

[32] D. M. Bossens and D. Tarapore. QED: using quality-environment-diversity to evolve resilient robot swarms. *IEEE Transactions on Evolutionary Computation*, 2020.

[33] A. Dharna, J. Togelius, and L. B. Soros. Co-generation of game levels and game-playing agents. In *Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment*, 2020.

[34] A. Dharna, A. K. Hoover, J. Togelius, and L. Soros. Transfer dynamics in emergent evolutionary curricula. *IEEE Transactions on Games*, 2022.

[35] M. Dennis, N. Jaques, E. Vinitsky, A. M. Bayen, S. Russell, A. Critch, and S. Levine. Emergent complexity and zero-shot transfer via unsupervised environment design. In *Advances in Neural Information Processing Systems 33*, 2020. URL https://proceedings.neurips.cc/paper/2020/hash/985e9a46e10005356bbaf194249f6856-Abstract.html.

[36] M. Jiang, E. Grefenstette, and T. Rocktäschel. Prioritized level replay. In *Proceedings of the 38th International Conference on Machine Learning, ICML*, 2021. URL http://proceedings.mlr.press/v139/jiang21b.html.

[37] M. Jiang, M. Dennis, J. Parker-Holder, J. N. Foerster, E. Grefenstette, and T. Rocktäschel. Replay-guided adversarial environment design. In *Advances in Neural Information Processing Systems 34*, 2021. URL https://proceedings.neurips.cc/paper/2021/hash/0e915db6326b6fb6a3c56546980a8c93-Abstract.html.

[38] J. Parker-Holder, M. Jiang, M. Dennis, M. Samvelyan, J. N. Foerster, E. Grefenstette, and T. Rocktäschel. Evolving curricula with regret-based environment design. *CoRR*, abs/2203.01302, 2022. doi:10.48550/arXiv.2203.01302.

[39] J. Mouret and J. Clune. Illuminating search spaces by mapping elites. *CoRR*, abs/1504.04909, 2015. URL http://arxiv.org/abs/1504.04909.

[40] M. C. Fontaine, J. Togelius, S. Nikolaidis, and A. K. Hoover. Covariance matrix adaptation for the rapid illumination of behavior space. In *Proceedings of the Genetic and Evolutionary Computation Conference*, 2020. doi:10.1145/3377930.3390232.

[41] M. C. Fontaine and S. Nikolaidis. Covariance matrix adaptation map-annealing. *arXiv preprint arXiv:2205.10752*, 2022.

[42] S. Pellegrinelli, H. Admoni, S. Javdani, and S. S. Srinivasa. Human-robot shared workspace collaboration via hindsight optimization. In *Proceedings of the International Conference on Intelligent Robots and Systems, IROS*, 2016. doi:10.1109/IROS.2016.7759147. URL https://doi.org/10.1109/IROS.2016.7759147.

[43] B. Tjanaka, M. C. Fontaine, J. Togelius, and S. Nikolaidis. Approximating gradients for differentiable quality diversity in reinforcement learning. *CoRR*, abs/2202.03666, 2022. URL https://arxiv.org/abs/2202.03666.

[44] T. Pierrot, V. Macé, F. Chalumeau, A. Flajolet, G. Cideron, K. Beguir, A. Cully, O. Sigaud, and N. Perrin-Gilbert. Diversity policy gradient for sample efficient quality-diversity optimization. In *Proceedings of the Genetic and Evolutionary Computation Conference*, pages 1075–1083, 2022.

[45] Y. Jiang, D. Salley, A. Sharma, G. Keenan, M. Mullin, and L. Cronin. An artificial intelligence enabled chemical synthesis robot for exploration and optimization of nanomaterials. *Science Advances*, 2022.

[46] A. Gaier, J. Stoddart, L. Villaggi, and P. J. Bentley. T-DominO: Exploring multiple criteria with quality-diversity and the tournament dominance objective. In *Proceedings of the 17th International Conference on Parallel Problem Solving from Nature*, 2022.

[47] A. Hagg, D. Wilde, A. Asteroth, and T. Bäck. Designing air flow with surrogate-assisted phenotypic niching. In *Proceedings of the International Conference on Parallel Problem Solving from Nature*, 2020.

[48] Y. Zhang, M. C. Fontaine, A. K. Hoover, and S. Nikolaidis. Deep surrogate assisted MAP-Elites for automated hearthstone deckbuilding. *CoRR*, abs/2112.03534, 2021. URL https://arxiv.org/abs/2112.03534.

[49] H. Zhang, M. Fontaine, A. Hoover, J. Togelius, B. Dilkina, and S. Nikolaidis. Video game level repair via mixed integer linear programming. In *Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment*, 2020.

[50] M. C. Fontaine, Y. Hsu, Y. Zhang, B. Tjanaka, and S. Nikolaidis. On the importance of environments in human-robot coordination. In *Robotics: Science and Systems*, 2021. doi:10.15607/RSS.2021.XVII.038.

[51] S. Javdani, S. S. Srinivasa, and J. A. Bagnell. Shared autonomy via hindsight optimization. In *Robotics: Science and Systems XI*, 2015. doi:10.15607/RSS.2015.XI.032. URL http://www.roboticsproceedings.org/rss11/p32.html.

[52] A. D. Dragan and S. S. Srinivasa. Formalizing assistive teleoperation. In *Robotics: Science and Systems VIII*, 2012. doi:10.15607/RSS.2012.VIII.010.

[53] S. Javdani, H. Admoni, S. Pellegrinelli, S. S. Srinivasa, and J. A. Bagnell. Shared autonomy via hindsight optimization for teleoperation and teaming. *The International Journal of Robotics Research*, 37(7):717–742, 2018.

[54] B. D. Ziebart, N. D. Ratliff, G. Gallagher, C. Mertz, K. M. Peterson, J. A. Bagnell, M. Hebert, A. K. Dey, and S. S. Srinivasa. Planning-based prediction for pedestrians. In *Proceedings of the International Conference on Intelligent Robots and Systems*, 2009. doi:10.1109/IROS.2009.5354147. URL https://doi.org/10.1109/IROS.2009.5354147.

[55] M. Liebhardt. OpenNI tracker. http://wiki.ros.org/openni_tracker, 2012.

[56] H. J. Jeon, D. P. Losey, and D. Sadigh. Shared autonomy with learned latent actions. *Robotics: Science and Systems (RSS)*, 2020.

[57] M. Chen, S. Nikolaidis, H. Soh, D. Hsu, and S. Srinivasa. Trust-aware decision making for human-robot collaboration: Model learning and planning. *ACM Transactions on Human-Robot Interaction (THRI)*, 9(2):1–23, 2020.

[58] R. Ye, W. Xu, H. Fu, R. K. Jenamani, V. Nguyen, C. Lu, K. Dimitropoulou, and T. Bhattacharjee. Rcare world: A human-centric simulation world for caregiving robots. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2022.

[59] Z. Erickson, V. Gangaram, A. Kapusta, C. K. Liu, and C. C. Kemp. Assistive gym: A physics simulation framework for assistive robotics. In *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, 2020.

[60] D. Das, S. Banerjee, and S. Chernova. Explainable AI for robot failures: Generating explanations that improve user assistance in fault recovery. In *Proceedings of the ACM/IEEE International Conference on Human-Robot Interaction*, 2021. doi:10.1145/3434073.3444657. URL https://doi.org/10.1145/3434073.3444657.

[61] B. Hayes and J. A. Shah. Improving robot controller transparency through autonomous policy explanation. In *Proceedings of the 2017 ACM/IEEE international conference on human-robot interaction*, pages 303–312, 2017.

[62] D. P. Kingma and J. Ba. Adam: A method for stochastic optimization. In *Proceedings of the 3rd International Conference on Learning Representations, ICLR*, 2015. URL http://arxiv.org/abs/1412.6980.

[63] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, et al. PyTorch: an imperative style, high-performance deep learning library. In *Advances in Neural Information Processing Systems*, 2019. URL https://proceedings.neurips.cc/paper/2019/file/bdbca288fee7f92f2bfa9f7012727740-Paper.pdf.

[64] IBM. ILOG CPLEX optimization studio. https://www.ibm.com/products/ilog-cplex-optimization-studio/cplex-optimizer, 2020.

[65] R. Diankov and J. Kuffner. OpenRAVE: A planning architecture for autonomous robotics. *Robotics Institute, Pittsburgh, PA, Tech. Rep. CMU-RI-TR-08-34*, 2008.

[66] S. Javdani. Shared autonomy via hindsight optimization code. https://github.com/personalrobotics/ada_assistance_policy, 2016.

[67] B. Tjanaka, M. C. Fontaine, D. H. Lee, T. T. M. Vu, Y. Zhang, S. Sommerer, N. Dennler, and S. Nikolaidis. pyribs: A bare-bones python library for quality diversity optimization. https://github.com/icaros-usc/pyribs, 2021.