2024 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)
October 14-18, 2024. Abu Dhabi, UAE

Roaming with Robots: Utilizing Artificial Curiosity in Global Path Planning for Autonomous Mobile Robots

Niklas Spielbauer[1] Till Laube[1] David Oberacker[1] Arne Roennau[1] Rüdiger Dillmann[1]

***Abstract*— Autonomous Mobile Robots are used with increasing frequency in inspection and maintenance tasks completing fixed goal sequences. The downtime robots experience between goals offers an opportunity to gather additional environment information instead of resting. Uncertainty in the amount of downtime available rules out the definition of a pre-determined schedule set by an external operator. Instead, the robot itself should decide dynamically, what information it should gather before its next task begins. This results in a multi-objective optimization problem trying to maximize information gain while utilizing as much of the available time as possible. We propose a genetic algorithm to solve the presented optimization problem and introduce two different models for artificial curiosity used inside the fitness function for gathering as much information as possible. For planning the genetic algorithm utilizes a multi-map approach using information and obstacle maps. We evaluated our models in a pre-defined and pre-mapped Gazebo environment with a given information map and evaluated their performance against an information-agnostic coverage algorithm. In this work, we show that utilizing artificial curiosity in path planning can result in major information gains by effectively using downtime.**

I. INTRODUCTION

Human curiosity is the main drive in experiencing new stimuli and finding novel information for all stages of human development. Curiosity-driven behavior leads to the discovery of new information about our environment, reducing uncertainty about assumptions up to uncovering the laws of nature themselves. Furthermore, it can lead to discovery and information gain even during mundane tasks.

Autonomous Mobile Robots (AMRs) have been replacing humans doing repetitive tasks in maintenance, inspection, logistics and other fields. During operation, robots rarely have 100% uptime and spend large amounts of time idle while waiting for the next task to begin. Utilizing this downtime more efficiently by e.g. collecting more information about the environment presents a large opportunity to make robots even more efficient.

Usually, downtime are not plannable thus making it infeasible to fill with predefined static tasks. Downtime information gathering should be flexible and autonomous, enabling the system itself to prioritize the information it wants to gather in a given amount of downtime.

Providing the robot itself with more agency will enhance the capabilities of autonomous robots and by extension robot teams used in approaches such as the one presented by Schnell et al. [1], reducing the workload for operators even further. Similarly, data collection approaches for active learning systems as presented by Puck et al. [2] can benefit from the system autonomously searching out potentially interesting data.

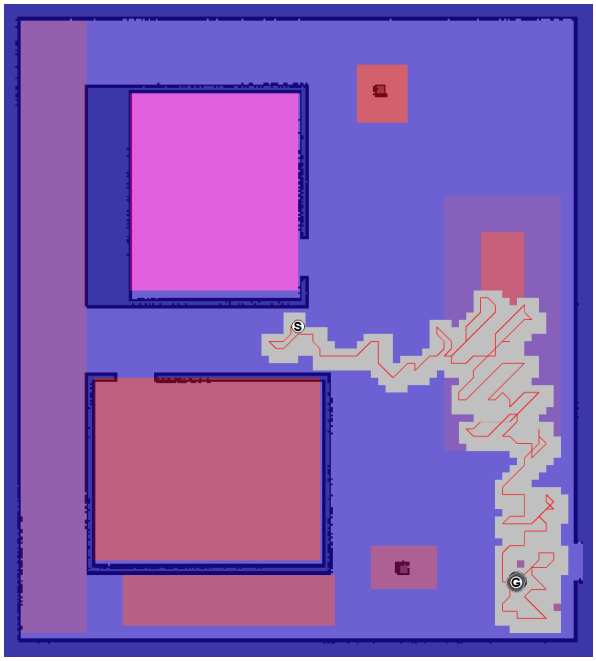


Fig. 1: Curious path planned through the used evaluation environment with pre-defined zones of interest. The zones of interest are shown in red and violet colors, depending on their information score. Obstacles in the environment are shown in dark blue. The start and goal points are marked with white circles while the area explored by the robot is marked in gray.

To ensure that tasks are still fulfilled with autonomously planned downtime while gathering as much information as possible, time and information-optimized path planning is needed, resulting in a multi-objective optimization problem. Liu et al. [3] propose a categorization of algorithms into classical, bionic and AI-based approaches. Within this differentiation, bionic approaches in particular were applied to solve multi-objective problems. In our work, we focused on genetic algorithms, a sub-group of bionic algorithms, allowing an in-depth understanding of the effect of the used models through the inspection of individual chromosomes.

[1] Department of Interactive Diagnosis and Service Systems (IDS), FZI Research Center for Information Technology, Haid-und-Neu-Straße 10–14, 76131 Karlsruhe, Germany

The research leading to these results has received funding in the FlexTools project under Grant Agreement No. 13IK032G by the Federal Ministry for Economic Affairs and Climate Action (BMWK).

Multi-Objective Genetic Algorithms (MOGAs) are well-explored in the field of global path planning in static environments. Chen et al. [4] employed a MOGA in path planning for cameras in scientific visualizations, optimizing for obstacle-free, short and smooth paths. An approach to pesticide application has been presented by Mahmud et al. [5], planning looped paths to apply pesticides to plants. Nazarahari et al. [6] propose the usage of an artificial potential field algorithm for the generation of initial paths which are in turn optimized by a MOGA inside a continuous environment for multi-robot applications focusing on path length, smoothness and safety. Li et al. [7] focused on the improvement of genetic operators while optimizing path length, path safety and energy consumption. Lamini et al. [8] presented an improved crossover method, generating multiple offspring and selecting the most promising ones.

Curiosity in both animals and humans has been investigated and classified extensively. Foundational work on curiosity has been done by Berlyne [9][10], categorizing curiosity into perceptual and epistemic behaviors. Perceptual curiosity, mainly observed in non-human animals and infants, is characterized by the driving force behind the motivation to seek out novel stimuli, while epistemic curiosity is a predominantly human trait, motivating the acquisition of knowledge that might be useful in the future. Additionally, Berlyne differentiates between specific curiosity, focused on a particular object, and diverse curiosity, a desire for cognitive stimulation. Kidd and Hayden [11] claim, that curiosity is linked to most human activities not directly related to basic survival, nourishment or reproduction, but a uniform definition of curiosity has not been archived.

Based on biological observations, artificial curiosity has been applied to both robotics and artificial intelligence applications. One of the first applications of artificial curiosity was presented by Schmidhuber [12] as two co-evolutionary models competing and cooperating to explore novel stimuli. Sun et al. [13] utilize reinforcement learning to optimize stimulus exposure time in a complementary arousal mechanism. Similarly, Fournier et al. [14] have an agent exploring a pick-and-place task autonomously while exploiting tutoring signals from human operators. Ramík et al. [15] used visual saliency to filter objects from an environment to model perceptual curiosity.

There are also several approaches using artificial curiosity in exploration. Pathak et al. [16] present an agent learning skills by examining its surroundings motivated by intrinsic rewards, defining curiosity as the error of prediction ability of the agent. Hutsebaut-Buysse [17] present a reinforcement learning-based approach of warehouse exploration based on implicitly learned world dynamics, whose learned policy does not need retraining in new environments and works with an information map to highlight interesting areas. Two different biologically inspired navigation methods were presented by Zhang et al. driving the agent to engage in activities based on remaining energy levels [18], and solving mazes through Q-Learning-based reinforcement learning [19]. Shi et al. [20] utilize intrinsic rewards for local planning, acknowledging the scarcity of extrinsic rewards. Zhelo et al. [21] utilize curiosity without a given occupancy map learning navigation policies by deep reinforcement learning.

Ngo et al. [22] utilize curiosity as intrinsic motivation to give rewards to a reinforcement learning model. Graziano et al. [23] use artificial curiosity to reduce the bandwidth needed for signal transmission in high-latency applications by discovering unobserved patterns in transmissions. Schaul et al. [24] define curiosity as the first derivative of the data compression ability of their current world model. With this definition, they search their data for new and learnable patterns that can in turn be used to enhance the world models' abilities.

While MOGAs are already well explored in current state-of-the-art, the topic of a planning algorithm for mobile robots allowing to plan both time and information optimized has not been explored. Such a planner is the first step to enable a robot to be curious and able to gather data without explicit orders, while adhering to external time constraints.

In this paper, we present an approach to utilize genetic algorithms to find paths for AMRs which are optimized for both time and information gathered by the robot. An example path can be seen in Figure 1. To make the planning algorithm aware of gathered information two different aspects of human curiosity are modeled inside the genetic algorithm. The curiosity models are evaluated against each other and a classical coverage planner.

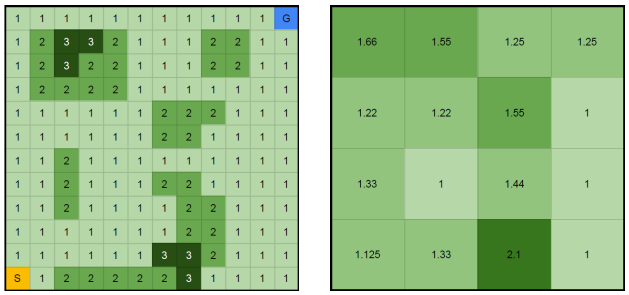
The planner is implemented as a Move-Base-Flex-based global planner [25], with evaluations taking place in a Gazebo environment.

The following paper is structured as follows. First, the genetic algorithm and the different curiosity models are introduced. Afterwards, the fine-tuning of the algorithm is presented and experiments and results are discussed, focusing on the information collected by the different curiosity models. Finally, the conclusion and outlook for future work are presented.

II. GENETIC ALGORITHM FOR TIME RESTRICTED CURIOUS PATH PLANNING

To ensure the robot can plan collision-free information-optimized paths we propose a multi-map setup. The primary map contains occupancy information of the robot's environment. Default 2D obstacle inflation is used to make sure every planned path on the obstacle map is traversable by the robot. The obstacle map is directly generated from the simulation environment, as we assume a completely static and known environment.

In addition to the obstacle map an information map is defined, marking interesting areas on the map. Each cell in the information map is assigned a cell information score I_{xy}. For now, we assume that the information score offers a comparable measure of the novelty of a cell while equal information scores also indicate a similar stimulus inside a static information environment. Both maps are discretized into grids with a cell width of 20cm as an input to the genetic planning algorithm. The information map is used to create



(a) Information map (b) Heuristic information map

Fig. 2: Information map and respective Heuristic Information map with a heuristic compression of $\frac{1}{4}$

an additional heuristic map, holding average information scores H_{xy} for larger areas on the map. An example of both information maps can be seen in Figure 2.

A. GENETIC ALGORITHM

Genetic algorithms are a well-explored approach to optimization problems in a variety of fields. In general, they are used to evolve a population P of possible problem solutions throughout multiple generations. In between generations the individuals, so-called chromosomes C are judged by a predefined fitness function. Promising individuals are selected from the population P_n and their genes are modified to create the population of the next generation P_{n+1}. Similar to Lamini et al. [8] and Li et al. [7] we assume a single chromosome to represent a possible path between two points on the map where every chromosome C_i consists of L_i cells as its genes g_i.

1) Initialization: To initialize the genetic algorithm an initial population P_0 is created. To increase the planner's efficiency only feasible paths can be part of the initial population, where a feasible path is a path where every selected cell is directly adjacent to its predecessor. Additionally, paths intersecting with obstacles are not considered feasible. The chromosome length of the individuals is randomly chosen in a range between the Manhattan distance from start to goal and the total number of cells on the map. The first gene in each chromosome is the starting point S. For every successor node g_i one of the eight neighboring nodes is randomly picked, as the next gene g_{i+1}, with all obstacle cells and cells outside the map boundaries being ignored. This procedure is repeated until the chromosome is filled with the number of genes L that was previously determined or until the goal state G is reached.

2) Fitness Evaluation: After the creation of each generation, all members of the population are assessed by a fitness function. The fitness function for our proposed algorithm is calculated as a weighted sum of individual fitness scores F_i, as seen in Figure 3.

$$F = \sum_{i=1}^{5} F_i \tag{1}$$

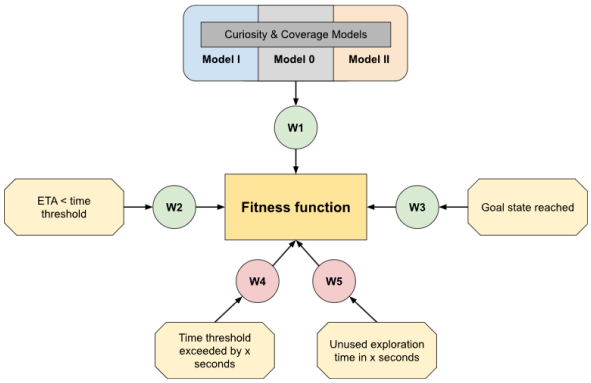


Fig. 3: Concept graphic of the fitness function used to judge chromosome fitness in our genetic algorithm, positive weights are shown in green, negative weights in red

For evaluating the time needed to traverse the path the traversal time t_{ETA} is estimated through its Angular and Cartesian distances.

$$t_{ETA} = v_{lin} \cdot \sum_{L} d_i + v_{ang} \cdot \sum_{L} \theta_i \tag{2}$$

Afterwards, rewards and penalties are calculated. Every individual C_i is rewarded for:

- Model information gathered during traversal of the path. This information score is dependent on the active curiosity model. Each gene g_n rewards the model with an information score M_n which is the sum of the information scores surrounding the cell represented by the gene.

$$F_1 = W_1 \cdot \sum_{n=0}^{L} M_n \text{ where } M_n = \sum_{x=\lceil -\frac{d}{2} \rceil}^{\lceil +\frac{d}{2} \rceil} \sum_{y=\lceil -\frac{d}{2} \rceil}^{\lceil +\frac{d}{2} \rceil} M_{\hat{x}+x\,\hat{y}+y} \tag{3}$$

- The calculated path duration t_{ETA} respecting the given time threshold t_{exec}.

$$F_2 = W_2 \cdot \begin{cases} 1 & \text{if } t_{ETA} < t_{exec} \\ 0 & \text{otherwise} \end{cases} \tag{4}$$

- The goal state G being reached with the last cell of the path represented by its last gene g_L.

$$F_3 = W_3 \cdot \begin{cases} 1 & \text{if } g_L = G \\ 0 & \text{otherwise} \end{cases} \tag{5}$$

Individuals are penalized for:

- Exceeding the given time threshold t_{exec}.

$$F_4 = -W_4 \cdot \max\left((t_{ETA} - t_{exec}), 0\right) \tag{6}$$

- Reaching the goal too early, leaving time unused.

$$F_5 = -W_5 \cdot \max\left((t_{exec} - t_{ETA}), 0\right) \tag{7}$$

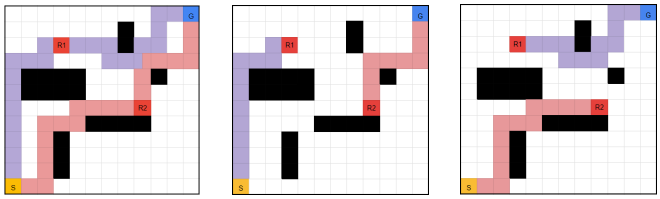
In general, the robot should receive a lower punishment for remaining unused time than for going over the ETA which

would violate external mission constraints. This means W_4 should always be larger than W_5. For all weights above it is assumed that $W_i > 0$.

3) Selection of Individuals: After evaluating the fitness of the individuals the selection for the next generation is performed using the layered method proposed by Li et al. [7] to ensure that only the traits of the best individuals are passed on.

4) Crossover and Mutation: After the selection, both crossover and mutation operations are performed on the new population. In general crossover operations combine the genes of two chromosomes to form two new chromosomes while mutation operations change a single chromosome randomly.

For crossover operations, the paths representing two different chromosomes are cut into two pieces each and recombined afterward. Cutting chromosomes at random genes and recombining them might result in unfeasible paths as the reconnection points might not overlap as shown in Figure 4. Reconnecting the loose ends of the chromosome parts directly might result in large additions of genes to the resulting paths that are not driven by optimization, leading to a potential drop in overall chromosome fitness. To ensure that only valid chromosomes are created with the crossover operation without the need for reconnections we limit crossover to only be performed if the two chromosomes share genes besides the start and end cells S and G, similar to the work by Li et al. [7].



(a) R1 and R2 are shown randomly selected genes where path 1 (purple) and path 2 (light red) will be split.

(b) Combining the first part of path 1 with the second part of path 2 results in a path that is not traversable.

(c) Combining the first part of path 2 with the second part of path 1 results in a path that is not traversable.

Fig. 4: Illustration of the crossover process when using random genes in the path. Visualizing how it creates potentially infeasible paths containing gaps between genes are generated.

Mutating genes to randomly selected cells will lead to a similar problem as the selected cell would need to be reconnected to its predecessor and successor. Li et al. [7] propose using an eight-cell neighborhood from the gene selected to be mutated to limit the needed reconnection lengths to a maximum of one cell. In case the mutated gene does create a gap in the path it is closed by calculating the position of the missing gene and filling the gap.

After crossover and mutation operations are performed the process is repeated until either a maximum number of generations N or a maximum planning time T_{plan} is reached.

B. CURIOSITY MODELS

As mentioned above, model information fitness F_1 is dependent on the active curiosity model. In this work, we present two approaches to partially model perceptive curiosity, the desire to experience new stimuli as well as the urge to explore heterogeneous areas. The models are inspired by the work of Berlyne [9][10] and Pathak [16]. Both curiosity models are compared to an information-agnostic coverage planner as a baseline. All models hold internal information maps, marking cells that have been visited by the robot as not containing more useful information.

1) Model 0: Simple Coverage Planner: As a baseline model, we introduce an information-agnostic model to the planner. As the planner is unable to calculate the information gain achieved with the exploration of different cells, it will maximize the covered area assuming homogeneous information density along the path. For Model 0 the agent gets rewarded with a constant information gain I_0 for every newly perceived cell on the map, independent of the actual information score of the cell.

$$M_{xy} = I_0 \cdot \begin{cases} 1 & \text{if cell is unexplored} \\ 0 & \text{otherwise} \end{cases} \tag{8}$$

2) Model I: Experience of New Stimuli: The first curiosity model Model I encourages the robot to explore cells that contain novel information. The degree of novelty experienced by the model is influenced by the number of similar stimuli experienced so far and the time interval between them. The more often a stimulus is experienced, the less interesting it becomes, permanently reducing its novelty. Experiencing a stimulus will also greatly reduce the model information gained by cells with a similar stimulus, which increases again with time passing between the experiences of a similar stimulus. The zones of interest defined in the detailed map are static, with the same stimuli being represented as equal information values I_{xy} as explained above. These effects are represented in Equation 9, with the model information score gained as M_{xy} and the time since the last similar stimulus was experienced as t_s and n_I the number of similar stimuli that have been experienced during the traversal of the path this far. n_I is tracked individually for every information value I_{xy}. w_t and w_n are weights to tune the novelty function properly.

$$M_{xy} = \max\left(I_{xy} \cdot \left(1 - \frac{w_t}{t_s + 1} - w_n \cdot (n_I - 1)\right), 0\right) \tag{9}$$

3) Model II: Exploration of Heterogeneous Areas: The second formulation of intrinsic motivation mentioned by Pathak et al. [16] is the drive that encourages the agent to perform actions that reduce its prediction uncertainty regarding the nodes on the map. As the prediction error cannot be calculated during planning, instead the heuristic map introduced earlier is used to encourage the agent to explore heterogeneous areas of the map, based on the assumption that these environments provide more new information and therefore stimulate a certain degree of curiosity.

For each cell, the heuristic error between the averaged and cell information is calculated representing the heterogeneity of the area.

$$M_{xy} = |H_{xy} - I_{xy}| \tag{10}$$

As the agent proceeds with evaluating the information score gathered along a path, both the heuristic and the detailed information density maps are updated. The detailed information density map is updated by setting the value of the cell the robot just explored as well as all other cells in its perception diameter d to zero. The heuristic map needs to be updated by recalculating the average of the compression areas, resulting in longer update runtimes. Runtime and heuristic error are both highly dependent on the heuristic compression value, as the heuristic compression is the inverse of the number of cells along one edge of the heuristic map e.g. a compression of 1/4 results in a heuristic map with 4x4 cells. As the heuristic map is updated every evaluation step, a lower heuristic compression value will result in larger compression areas and larger computations, while a higher heuristic compression value will result in smaller computations. As this update needs to be done for all paths in every generation, increasing the computational cost of calculating one heuristic error by a small amount can have severe consequences for the overall runtime.

Choosing the right heuristic compression is not trivial. Larger heuristic grid sizes will average over a larger number of cells, leading to similar heuristic scores across the map, but potentially large heuristic errors inside the individual cells. If the grid size is chosen too large the amount of different zones will be too small, resulting in the robot not exploring heterogeneous zones, but areas where the information score difference to the global average is the largest. Smaller grid sizes will on average have lower heuristic errors, as the map gets more and more similar to the detailed map the smaller the grid size is.

III. RESULTS AND EXPERIMENTS

With our presented approach we want to improve the information gathered while traversing a map towards a goal. For model evaluation, we consider both the amount of unbiased information U_{ges} gathered by the robot while traversing the path as well as the information density D defined as

$$U_{ges} = \sum^{L} U_i \text{ where } U_i = \sum_{x=\lceil -\frac{d}{2} \rceil}^{\lceil +\frac{d}{2} \rceil} \sum_{y=\lceil -\frac{d}{2} \rceil}^{\lceil +\frac{d}{2} \rceil} I_{\hat{x}+x\,\hat{y}+y} \tag{11}$$

$$D = \frac{U_{ges}}{L} \tag{12}$$

We expect that the curious models will achieve higher unbiased information scores U_{ges} and density D than the coverage path planner while the goal state G is reached in the given time.

We evaluate our models in a static, fully known environment with given interest zones defined by constant information values. The robot is given a fixed time limit for path planning and execution. We chose a limit of $T = 240s$ to make sure meaningful optimizations can be done in both steps. The environment is split into a grid with a grid size of 20cm, where we assume the robot's perception diameter to be $d = 3$ meaning the robot is capable of perceiving a 16-cell neighborhood around itself. For all evaluation procedures, there are 20 paths simulated to each of the two goal points shown in Figure 5. The goal points were chosen to evaluate performance with a target in a homogeneous area (Goal 1: $(-2|4)$) as well as a heterogeneous area (Goal 2: $(6|-7)$).

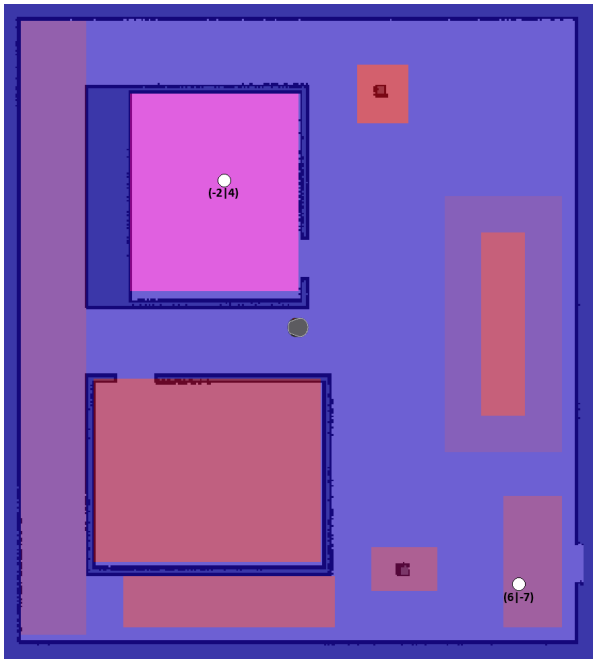


Fig. 5: The two goal points used for the evaluation of the presented curiosity models

A. INITIAL TUNING OF GA PARAMETERS

Before the evaluation of the individual models, some general parameters need to be adjusted, to ensure good baseline performance of the algorithm.

To determine the final parameters shown in Table I the following procedures were taken. First basic GA parameters, such as crossover/mutation split and mutation ratio per generation were tuned to generate promising paths maximizing the chance of successfully reaching the goal. To ensure that time constraints are adhered to, the ETA reward W_2 was increased as it proved to be more effective than increasing the overtime penalty $W4$. Model Reward W_1, Goal Reward W_3, and Inactivity Penalty W_5 were also tuned to ensure the model meets the given time constraints. Additionally, the path duration heuristic was adjusted to fit the actual robot travel times by adjusting both approximated velocities v_{lin} and v_{ang}. For Model I both decay weights w_t and w_n were tuned to fit the information scores in the given information map which were defined as $0 < I_{xy} < 100$.

The heuristic compression of the map for the second model was determined by comparing the unbiased information U_{ges},

TABLE I: Adjustable parameters for the Multi-Objective Genetic Algorithm and their final values

Parameter	Symbol	Value	Description
Population Size	N_p	200	Number of chromosomes per population
Max. Generations	N	2000	Max. number of generations iterated over
Crossover Ratio	r_c	0.7	Percentage of chromosomes that are crossed
Mutation Ratio	r_m	0.3	Percentage of chromosomes that are mutated
Mutation Percentage	p_m	0.1	Percentage of genes mutated in each chromosome
Lin. Velocity	v_{lin}	2.0	Robots average linear velocity (m/s)
Ang. Velocity	v_{ang}	1.0	Robots average angular velocity (rad/s)
Model Reward	W_1	3	Curiosity model's influence on the fitness function
ETA Reward	W_2	500000	Single time reward when travel time is below limit
Goal Reward	W_3	1000000	Single time reward when goal is reached
Overtime Penalty	W_4	-200	Penalty per second of overtime
Inactivity Penalty	W_5	-10	Penalty per second of unused downtime
Stimulus Time Decay	w_t	0.4	Decay weight for time between similar stimuli
Stimulus Amount Decay	w_n	0.01	Decay weight for number of similar stimuli
Heuristic Compression	c_h	0.25	Compression of the heuristic information map

information density D and the maximum generation reached at heuristic compressions c_h of 0.5, 0.25, 0.125 and 0.0625 for a maximum of 100 generations and $t_{plan} = 120s$. Except for $c_h = 0.5$, all compressions reached the maximum number of generations. Moving forward a compression of $c_h = 0.25$ was used, as it provided both the best unbiased information and information density scores.

B. MODEL EVALUATION

The total time available to the model is given as $T = t_{plan} + t_{exec}$ and while it can be assumed that longer planning times will result in more optimal paths, reducing the execution time will result in shorter paths, limiting the possible information gain.

In our experiments, we focused on t_{plan}/t_{exec} ratios of 20/80, 40/60 and 60/40 as early experiments clearly showed a large drop in performance for a ratio of 80/20. Figures 6 and 7 show the unbiased information values gathered for all three models and both goals. All configurations where the median for the estimated path time t_{ETA} exceeded the travel time limit are greyed out as they show that the paths are not adhering to the given constraints. Those invalid configurations can show higher information gain than other, valid, configurations as they ignore the implicit maximum path length limit given by t_{exec}. In general, the unbiased information values for Goal 2 are lower than the ones for Goal 1 as the latter lies inside a homogeneous area of interest with high information scores I_{xy}, resulting in large amounts of information gathered in that area.

While staying within the given time constraints Model 0 performs better for both goals the more time it has to traverse the path and gather information. On average the configurations for Model 0 broke the constraints only once, for the 60/40 ratio for Goal 2, while the 80/20 ratio performed best for both Goals.

Model I shows a similar behavior for Goal 1, gathering more information with higher execution times. For Goal 2 Model I shows better performance when having more time to plan, but drops off when t_{exec} and with that viable paths get too short. Model I adheres to the constraints the best, with not a single configuration expected to execute the path outside

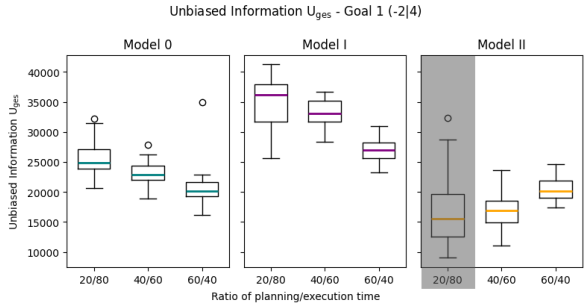


Fig. 6: Unbiased information gained by all models for Goal $(-2|4)$, configurations that are not expected to reach the goal in time are grayed out

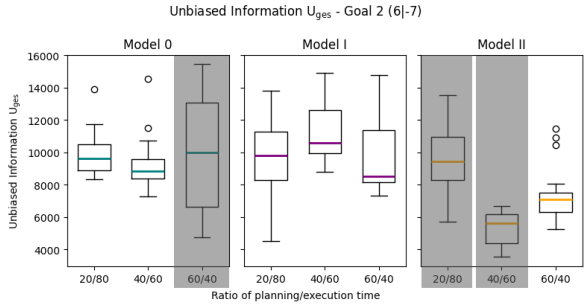


Fig. 7: Unbiased information gained by all models for Goal $(6|-7)$, configurations that are not expected to reach the goal in time are grayed out

the given time frame. While for Goal I the 20/80 ratio shows better performance than the 40/60 ratio the latter shows a lower spread in information gathered while maintaining high performance for both goals, making it the more promising configuration.

Model II shows increasing performance with larger planning times for Goal 1, while only the 60/40 ratio was expected to find a valid path inside the execution time for Goal 2. Model II struggled the most with adhering to the external constraints, with 1/3 configurations for Goal 1 and 2/3 configurations for Goal 2 expected to exceed the given

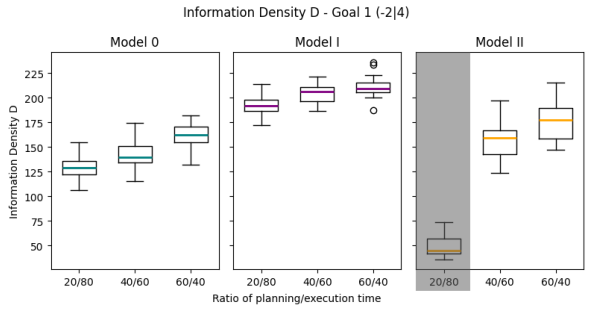


Fig. 8: Average Information Density gained by all models for Goal $(-2|4)$, configurations that are not expected to reach the goal in time are grayed out

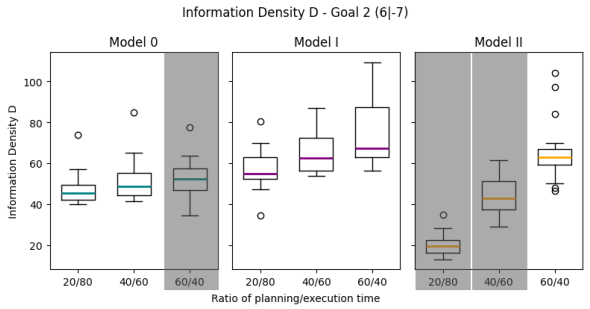


Fig. 9: Average Information Density gained by all models for Goal $(6|-7)$, configurations that are not expected to reach the goal in time are grayed out

time limit. For Goal 1 the best configuration was shown to be the 60/40 ratio, which is also the only one adhering to the constraints for Goal 2.

In general Model I shows the highest information scores for both Goals, followed by Model 0 except the 60/40 ratio for Goal 1, where Model II matches the performance of Model 0. When comparing the best-found configurations Model I with a 40/60 ratio was found to gather a median information increase of +10.29% to +33.61% with the mean increasing by +14.37% to +29.14% compared to Model 0 with a 20/80 ratio depending on the goal point. Model II performs worse than the baseline model showing a median information decrease of -18.96% to -26.30% with the mean changing by -20.5% to -25.4% compared to the best Model 0 configuration.

The information density for all three models and the two different goal points is plotted in Figures 8 and 9. When comparing the information density for all three models and both goal points a clear trend across all models is visible. The more time is allocated for planning and therefore optimizing the path, the higher the information density is. For Goal 1 Model I again shows the highest performance across configurations, while for Goal 2 the only valid configuration for Model II outperforms Model 0 showing a similar performance to Model I. In general, configurations that violate the constraints show low information density scores, as their additional information gain is largely based on the generation of longer paths. While the increasing information density with more limited execution time proves the optimization ability of the genetic algorithm itself regarding the information gain maximization within a given time frame a path with optimized information efficiency is not necessarily the most informative path to be found in the given time T.

All in all Model I provides a well-working implementation of perceptual curiosity using novelty that encourages the exploration of new states, confirming the effectiveness of artificial curiosity utilizing areas of interest similar to Hutsebaut-Buysse [17].

Model II does in general perform worse than a simple coverage planner. When comparing the number of generations each model was iterating over in the given timeframe as shown in Figure 10 it is obvious, that the performance of Model II was severely encumbered by runtime issues caused by updating the heuristic map, hindering its ability to generate optimized paths. As can be seen in Figure 9, Model II does show higher information density than Model 0, promising better results if the model was able to iterate over more generations. A lower heuristic compression could be used to improve the update speed but could result in a reduction in heuristic errors and decrease the optimization potential of the model. More optimization in calculating the heuristic map needs to be done to calculate heuristic errors more efficiently, decreasing the high computational cost of recalculations between generations. In general, while the results for both models are promising, with Model I strictly outperforming the coverage planner, the computational costs of Multi-Objective Genetic Algorithms need to be considered while pursuing this approach to make it applicable in 3D or on larger maps. Still, a novelty-based implementation of perceptual curiosity has been shown to improve the amount of gathered information, resulting in effective downtime minimization in the context of path planning.

IV. CONCLUSIONS AND FUTURE WORKS

In this work, we presented two separate models for utilizing perceptual curiosity inside a Multi-Objective Genetic Algorithm to maximize information gained in time-constrained path planning. Both models utilize a multi-map setup to maximize information gained while traversing a path towards a goal. The models were evaluated separately and compared with an information-agnostic coverage planner.

While the model based on a heuristic error showed diminished performance due to computational costs multiple configurations of the Multi-Objective Genetic Algorithm were shown to keep the given time constraints while gathering more information than the reference model. A second novelty-based model outperformed the reference model in all configurations gathering 10.29%-33.61% more information on average. The information density of paths generated by both curiosity models was also shown to be on average higher than the density of paths generated by the reference model. With the presented models we introduced two promising

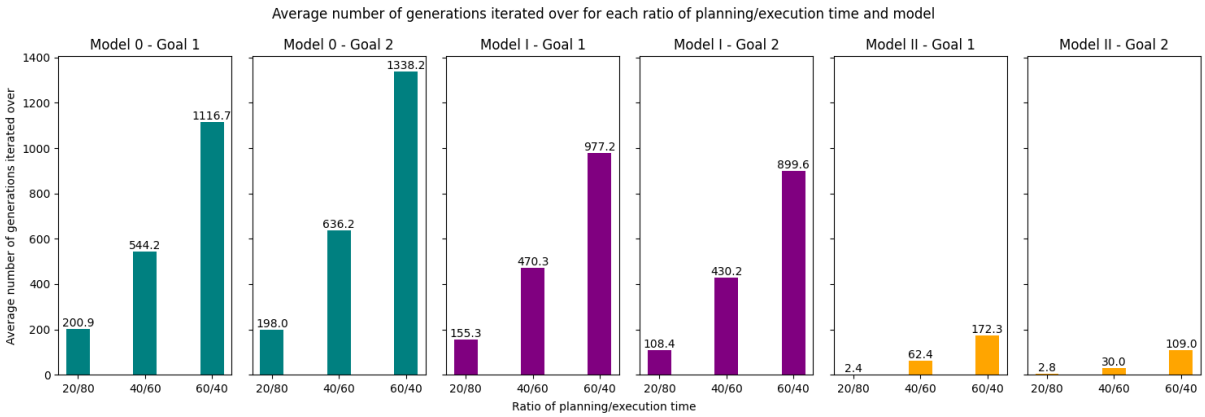


Fig. 10: Overview of the average number of generations iterated over for each model and ratio

implementations of parts of artificial perceptual curiosity. Utilizing the models to give the robots more agency we can increase the potential of AMRs used in logistics and inspection operations, allowing for efficient downtime usage. In the future, the presented approach should be expanded to allow for the maps to be time-variable, allowing for the use of dynamic environment models while planning. While the presented curiosity models showed promising results, the model selection should be expanded, also encompassing other parts of perceptual curiosity. Additionally, the approach should be transferred to a 3D map, as 2D maps are not sufficient to model the information found in a 3D environment.

References

[1] T. Schnell, D. Oberacker, F. Exner, L. Puck, M. G. Besselmann, N. Spielbauer, C. Plasberg, A. Roennau, and R. Dillmann, "An efficient scalable autonomy approach for teams of heterogeneous mobile robots," in *2023 IEEE 19th International Conference on Automation Science and Engineering (CASE)*. IEEE, 2023, pp. 1–7.

[2] L. Puck, N. Spielbauer, D. Schäble, T. Schnell, T. Büttner, A. Roennau, and R. Dillmann, "Distributed active learning for semantic segmentation on walking robots," in *2021 20th International Conference on Advanced Robotics (ICAR)*. IEEE, 2021, pp. 461–467.

[3] L. Liu, X. Wang, X. Yang, H. Liu, J. Li, and P. Wang, "Path planning techniques for mobile robots: Review and prospect," *Expert Systems with Applications*, p. 120254, 2023.

[4] Z. Chen, J. Zhou, R. Sun, and L. Kang, "A new evolving mechanism of genetic algorithm for multi-constraint intelligent camera path planning," *Soft Computing*, vol. 25, pp. 5073–5092, 2021.

[5] M. S. A. Mahmud, M. S. Z. Abidin, Z. Mohamed, M. K. I. Abd Rahman, and M. Iida, "Multi-objective path planner for an agricultural mobile robot in a virtual greenhouse environment," *Computers and electronics in agriculture*, vol. 157, pp. 488–499, 2019.

[6] M. Nazarahari, E. Khanmirza, and S. Doostie, "Multi-objective multi-robot path planning in continuous environment using an enhanced genetic algorithm," *Expert Systems with Applications*, vol. 115, pp. 106–120, 2019.

[7] K. Li, Q. Hu, and J. Liu, "Path planning of mobile robot based on improved multiobjective genetic algorithm," *Wireless Communications and Mobile Computing*, vol. 2021, pp. 1–12, 2021.

[8] C. Lamini, S. Benhlima, and A. Elbekri, "Genetic algorithm based approach for autonomous mobile robot path planning," *Procedia Computer Science*, vol. 127, pp. 180–189, 2018.

[9] D. E. Berlyne *et al.*, "A theory of human curiosity," 1954.

[10] D. E. Berlyne, "Curiosity and exploration: Animals spend much of their time seeking stimuli whose significance raises problems for psychology." *Science*, vol. 153, no. 3731, pp. 25–33, 1966.

[11] C. Kidd and B. Y. Hayden, "The psychology and neuroscience of curiosity," *Neuron*, vol. 88, no. 3, pp. 449–460, 2015.

[12] J. Schmidhuber, "Artificial curiosity based on discovering novel algorithmic predictability through coevolution," in *Proceedings of the 1999 congress on evolutionary computation-cec99 (cat. no. 99th8406)*, vol. 3. IEEE, 1999, pp. 1612–1618.

[13] C. Sun, H. Qian, and C. Miao, "From psychological curiosity to artificial curiosity: Curiosity-driven learning in artificial intelligence tasks," *arXiv preprint arXiv:2201.08300*, 2022.

[14] P. Fournier, O. Sigaud, and M. Chetouani, "Combining artificial curiosity and tutor guidance for environment exploration," in *Workshop on Behavior Adaptation, Interaction and Learning for Assistive Robotics at IEEE RO-MAN 2017*, 2017.

[15] D. M. Ramík, C. Sabourin, and K. Madani, "Autonomous knowledge acquisition based on artificial curiosity: Application to mobile robots in an indoor environment," *Robotics and Autonomous Systems*, vol. 61, no. 12, pp. 1680–1695, 2013.

[16] D. Pathak, P. Agrawal, A. A. Efros, and T. Darrell, "Curiosity-driven exploration by self-supervised prediction," in *International conference on machine learning*. PMLR, 2017, pp. 2778–2787.

[17] M. Hutsebaut-Buysse, F. G. Guinjoan, E. Rademakers, S. Latré, A. B. Temsamani, K. Mets, E. Mannens, and T. De Schepper, "Directed real-world learned exploration," in *2023 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2023, pp. 5227–5234.

[18] J. Zhang, X. Ruan, J. Huang, J. Chai, and Y. Wu, "A curiosity-based mobile robot path planning method," in *2019 IEEE 4th Advanced Information Technology, Electronic and Automation Control Conference (IAEAC)*, vol. 1. IEEE, 2019, pp. 476–480.

[19] X. Zhang, Y. Liu, D. Hu, and L. Liu, "A maze robot autonomous navigation method based on curiosity and reinforcement learning," in *The 7th Int. Workshop on Advanced Computational Intelligence and Intelligent Informatics (IWACIII 2021), Article*, no. M1-6, 2021, p. 1.

[20] H. Shi, L. Shi, M. Xu, and K.-S. Hwang, "End-to-end navigation strategy with deep reinforcement learning for mobile robots," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 4, pp. 2393–2402, 2019.

[21] O. Zhelo, J. Zhang, L. Tai, M. Liu, and W. Burgard, "Curiosity-driven exploration for mapless navigation with deep reinforcement learning," *arXiv preprint arXiv:1804.00456*, 2018.

[22] H. Ngo, M. Luciw, A. Forster, and J. Schmidhuber, "Learning skills from play: artificial curiosity on a katana robot arm," in *The 2012 international joint conference on neural networks (IJCNN)*. IEEE, 2012, pp. 1–8.

[23] V. Graziano, T. Glasmachers, T. Schaul, L. Pape, G. Cuccu, J. Leitner, and J. Schmidhuber, "Artificial curiosity for autonomous space exploration," *Acta Futura*, vol. 4, pp. 41–51, 2011.

[24] T. Schaul, Y. Sun, D. Wierstra, F. Gomez, and J. Schmidhuber, "Curiosity-driven optimization," in *2011 IEEE Congress of Evolutionary Computation (CEC)*. IEEE, 2011, pp. 1343–1349.

[25] S. Pütz, J. Santos Simón, and J. Hertzberg, "Move base flex a highly flexible navigation framework for mobile robots," in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2018, pp. 3416–3421.