

NNgTL: Neural Network Guided Optimal Temporal Logic Task Planning for Mobile Robots

Ruijia Liu, Shaoyuan Li and Xiang Yin

Abstract—In this work, we investigate task planning for mobile robots under linear temporal logic (LTL) specifications. This problem is particularly challenging when robots navigate in continuous workspaces due to the high computational complexity involved. Sampling-based methods have emerged as a promising avenue for addressing this challenge by incrementally constructing random trees, thereby sidestepping the need to explicitly explore the entire state-space. However, the performance of this sampling-based approach hinges crucially on the chosen sampling strategy, and a well-informed heuristic can notably enhance sample efficiency. In this work, we propose a novel *neural-network guided* (NN-guided) sampling strategy tailored for LTL planning. Specifically, we employ a multi-modal neural network capable of extracting features concurrently from both the workspace and the Büchi automaton. This neural network generates predictions that serve as guidance for random tree construction, directing the sampling process toward more optimal directions. Through numerical experiments, we compare our approach with existing methods and demonstrate its superior efficiency, requiring less than 15% of the time of the existing methods to find a feasible solution.

I. INTRODUCTION

With the ongoing development and widespread deployment of mobile robots, there has been an increasing focus on path planning for high-level tasks. Among the formal languages used for specifying such complex tasks, Linear Temporal Logic (LTL) stands out as a widely adopted choice. LTL provides a structured means for the users to articulate complex requirements, such as navigating a robot to a target region without collision with obstacles or ensuring specific locations are visited infinitely often [1]–[5]. In recent years, the field of robot path planning for LTL tasks has witnessed extensive investigation, spurred by its broad applications. These applications encompass a wide range of scenarios, such as environmental surveillance [6], search and rescue missions [7], and intelligent warehousing systems [8], [9].

In the context of LTL path planning, one of the most foundational methods is the automata-theoretical approach based on finite abstractions [6], [10]–[15]. This approach involves the creation of a discrete abstraction of the workspace, which effectively captures the robot's mobility constraints. Subsequently, the discrete abstraction is synchronized with an automaton representation of the LTL task. This synchronization enables the formulation of the planning problem as a graph-search problem within the product space. However, this graph-search approach, although conceptually powerful,

faces a significant computational challenge. In particular, as the system's dimensionality increases, the state-space of the finite abstraction grows exponentially, rendering the graph-search problem infeasible.

Sampling-based methods, such as rapid random trees (RRT), have emerged as a promising solution to tackle the computational challenges associated with path planning in continuous state-spaces [16]. More recently, sampling-based algorithms have been introduced to enhance the computational efficiency of solving LTL planning problems [17]–[21]. For example, in [19], a sampling-based algorithm, inspired by the RRT*, is proposed to find optimal paths that satisfy LTL tasks with a prefix-suffix structure. This algorithm circumvents the necessity of explicitly exploring the entire state-space by incrementally constructing random trees over the product state-space. Building upon this advancement, [20] further introduces a biased sampling strategy that leverages automata information to significantly improve sample efficiency. Moreover, in [21], the authors extend the methods in [20] to continuous workspaces without discretization. Specifically, they introduce an abstraction-free planning algorithm that directly samples within a continuous space and integrates these samples with automata states.

In sampling-based LTL planning, one of the key factor is how each new state is sampled. While biased sampling strategies offer an enhanced approach compared to uniform sampling, they primarily rely on distance information within the Büchi automaton. This approach neglects valuable insights from the workspace, such as the physical feasibility of task progression. However, integrating continuous workspace information with the Büchi automaton is very challenging due to the inherently heterogeneous structures of these components. Recently, there has been a notable progress in leveraging the power of neural networks for expediting solutions to intricate problems. For instance, within the domain of path planning, neural networks have been employed to predict the probability distribution of optimal paths [22]–[25]. Moreover, neural networks have been applied, in an end-to-end fashion, for generating solutions to temporal logic tasks [26]–[29].

In this paper, we introduce a novel neural-network guided (NN-guided) sampling strategy designed specifically for LTL planning. Our approach builds upon the basic architecture of the sampling-based algorithm for continuous workspaces, as established in [21]. However, instead of relying solely on automata information, we incorporate a multi-modal neural network capable of jointly extracting features from both the workspace and the Büchi automaton. This neural network offers predictive capabilities to steer the sampling

This work was supported by the National Natural Science Foundation of China (62173226, 62061136004).

R. Liu, S. Li and X. Yin are with the Department of Automation, Shanghai Jiao Tong University, Shanghai 200240, China. E-mail: {liuruijia, syli, yinxiang}@sjtu.edu.cn

process in directions that are more likely to yield optimal solutions, ones that are not only task-progressive but also feasible within the workspace. To demonstrate the efficiency of our approach, we compare the proposed NN-guided sampling strategy with existing sampling strategies on a set of randomly generated problem instances. The statistical results underscore the effectiveness of our new strategy, as it requires less than 15% of the time needed by the existing strategies to find a feasible solution.

II. PROBLEM FORMULATION

A. System Model

We consider the scenario, where a single robot navigating in a two-dimensional continuous workspace represented by a compact subset $\mathcal{W} \subseteq \mathbb{R}^2$. The workspace is assumed to be partitioned as $\mathcal{W} = \mathcal{W}_{\text{free}} \dot{\cup} \mathcal{O}$, where \mathcal{O} is an open subset representing obstacle regions and $\mathcal{W}_{\text{free}}$ is the free region in which the robot can move. The free workspace is further partitioned as m labeled regions of interest and a non-labeled region $\mathcal{W}_{\text{free}} = \mathcal{R}_1 \dot{\cup} \dots \dot{\cup} \mathcal{R}_m \dot{\cup} \mathcal{R}_{\text{non}}$.

The mobility of the robot can be captured by a weighted transition system

$$TS = (\mathcal{W}, \mathbf{x}_0, \rightarrow_T, \mathcal{AP}, L, C),$$

where \mathcal{W} is set of infinite positions of the robot; \mathbf{x}_0 is its initial position; $\rightarrow_T \subseteq \mathcal{W} \times \mathcal{W}$ is the transition relation such that: for any $\mathbf{x}, \mathbf{x}' \in \mathcal{W}$, we have $(\mathbf{x}, \mathbf{x}') \in \rightarrow_T$ if along the straight line between \mathbf{x} and \mathbf{x}' (i) does not intersect with \mathcal{O} ; and (ii) crosses any boundary of labeled regions at most once [21]; $\mathcal{AP} = \{\pi^1, \dots, \pi^m\}$ is the set of atomic propositions with labeling function $L : \mathcal{W} \rightarrow \mathcal{AP}$ such that $L(\mathbf{x}) = \pi^i$ iff $\mathbf{x} \in \mathcal{R}_i$; $C : \mathcal{W} \times \mathcal{W} \rightarrow \mathbb{R}^+$ is the weight function calculating the Euclidean distance, i.e., $C(\mathbf{x}_1, \mathbf{x}_2) = \|\mathbf{x}_1 - \mathbf{x}_2\|_2$.

An infinite *path* is a sequence $\tau = \tau(1)\tau(2) \dots \in \mathcal{W}^\omega$ such that $\tau(1) = \mathbf{x}_0$ and $(\tau(k), \tau(k+1)) \in \rightarrow_T, k = 1, 2, \dots$. The *trace* of infinite path τ is $\text{trace}(\tau) = L(\tau(1))L(\tau(2)) \dots \in (\mathcal{AP})^\omega$. We say a path τ is finite if it is a finite sequence of points, and we denote by $|\tau|$ its *length*. For a finite path τ , its cost $J(\tau)$ is defined as the cumulative distances between consecutive states, i.e., $J(\tau) = \sum_{k=1}^{|\tau|-1} C(\tau(k), \tau(k+1))$.

B. Temporal Logic Tasks

The formal specification of robot is captured by an linear temporal logic formula without next (LTL_○), which is widely used in robot task planning in continuous workspace [30]. The syntax of LTL_○ is given as follows:

$$\phi ::= \text{true} \mid \pi \in \mathcal{AP} \mid \neg \phi \mid \phi_1 \wedge \phi_2 \mid \phi_1 \mathcal{U} \phi_2,$$

where \neg and \wedge are Boolean operators “negation” and “conjunction”, respectively; \mathcal{U} is the temporal operator “until”, which can further induce temporal operators “eventually” \Diamond and “always” \Box .

An LTL formula ϕ is evaluated over an infinite word $\sigma = \sigma(1)\sigma(2) \dots \in \mathcal{AP}^\omega$. We denote by $\sigma \models \phi$ that word σ satisfies LTL formula ϕ ; the reader is referred to [31] for the formal semantics. We denote by $\text{Words}(\phi)$ the set of all words satisfying formula ϕ . It is well-known that,

for any LTL formula ϕ , $\text{Words}(\phi)$ can be accepted by a non-deterministic Büchi automaton (NBA) [31]. Formally, an NBA is a tuple $B = (\mathcal{Q}_B, \mathcal{Q}_B^0, \Sigma, \rightarrow_B, \mathcal{Q}_B^F)$, where \mathcal{Q}_B is the set of states, \mathcal{Q}_B^0 is the set of initial states, $\Sigma = \mathcal{AP}$ is the alphabet, $\rightarrow_B \subseteq \mathcal{Q}_B \times \Sigma \times \mathcal{Q}_B$ is the transition relation, and \mathcal{Q}_B^F is the set of accepting states. For simplicity, we assume that the initial-state is unique, i.e., $\mathcal{Q}_B^0 = \{q_B^0\}$.

An infinite run ρ_B of B over an infinite word $\sigma = \pi_0\pi_1 \dots \in (\mathcal{AP})^\omega$ is a sequence $\rho_B = q_B^0 q_B^1 q_B^2 \dots$ such that $q_B^0 \in \mathcal{Q}_B^0, (q_B^i, \pi_i, q_B^{i+1}) \in \rightarrow_B, \forall i = 0, 1, \dots$. We say infinite run ρ_B is accepting if it contains accepting states infinite number of times; we say a word is accepting if it induces an accepting run, and we denote by \mathcal{L}_B the set of all accepting words of B . Hereafter, B is referred to as the NBA that accepts ϕ , i.e., $\mathcal{L}_B = \text{Words}(\phi)$.

C. LTL Task Planning Problem

To fulfill an LTL_○ formula ϕ , the robot needs to execute an infinite path. For the purpose of planning, it suffices to consider paths in the “prefix-suffix” structure $\tau = \tau^{\text{pre}}[\tau^{\text{suf}}]^\omega$, where the prefix part τ^{pre} repeats only once and the suffix part τ^{suf} repeats indefinitely [10]. The cost of prefix-suffix path τ is defined by $J(\tau) = \lambda J(\tau^{\text{pre}}) + (1 - \lambda)J(\tau^{\text{suf}})$, where $\lambda \in [0, 1]$ is a user-defined weight coefficient.

Then our objective is to find a plan for the robot in the prefix-suffix structure such that a given LTL formula is fulfilled with minimum cost.

Problem 1: Given LTL_○ formula ϕ , determine a prefix-suffix path τ in transition system TS such that (i) $\text{trace}(\tau) \models \phi$; and (ii) for any prefix-suffix path τ' such that $\text{trace}(\tau') \models \phi$, we have $J(\tau) \leq J(\tau')$.

III. SAMPLING-BASED TASK PLANNING

To solve Problem 1, a typical approach is to perform graph-search on the *product* of TS and B [6]. However, when the state space \mathcal{W} is continuous, building the entire product space is infeasible even by discretizations. Therefore, in [21], the authors proposed a (continuous space) sampling-based algorithm, called TL-RRT*, that incrementally builds trees on-the-fly without construct the entire product space a priori. Our work is build upon the sampling-based approach; therefore, we review it briefly in this section. The readers are referred to [21] for more details on this method.

A. Main Sampling-Based Algorithm

The TL-RRT* algorithm is essentially a random search over the product space $\mathcal{Q}_P := \mathcal{W}_{\text{free}} \times \mathcal{Q}_B$. It consists of two parts, prefix part and suffix part, and for each part, a similar random tree construction is employed in order to search the optimal path. We briefly sketch the procedures.

1) *Prefix Search:* Starting from the initial state $q_0 = (\mathbf{x}_0, q_B^0)$, one first builds a prefix tree $\mathcal{T} = (\mathcal{V}_{\mathcal{T}}, \mathcal{E}_{\mathcal{T}}, \text{Cost})$ by incrementally adding vertices, where $\mathcal{V}_{\mathcal{T}}$ and $\mathcal{E}_{\mathcal{T}}$ are the sets of vertices and edges in \mathcal{T} , respectively. The tree contains some *goal states* defined by $\mathcal{Q}_{\text{goal}} := \mathcal{W}_{\text{free}} \times \mathcal{Q}_B^F$. By projecting paths in tree onto \mathcal{W} , the path from the initial

state q_0 to each goal state $q_{\text{goal}} \in \mathcal{V}_{\mathcal{T}} \cap \mathcal{Q}_{\text{goal}}$ forms a prefix path τ^{pre} , whose cost is stored by function $\text{Cost} : \mathcal{V}_{\mathcal{T}} \rightarrow \mathbb{R}^+$.

2) *Suffix Search*: The suffix part is very similar to the prefix part, the main difference is that one needs to construct a set of random trees rooted from goal states $\mathcal{P} := \mathcal{T} \cap \mathcal{Q}_{\text{goal}}$ in the prefix tree such that they can return back to the goal states. This gives us a suffix plan τ^{suf} from each goal state, and the prefix-suffix plan with minimum cost among all goal states is the final optimal plan.

3) *Constructing Random Trees*: In the above two parts, the key is construction of random tree $\mathcal{T} = (\mathcal{V}_{\mathcal{T}}, \mathcal{E}_{\mathcal{T}}, \text{Cost})$ on-the-fly. The main construction steps are as follows.

- 1) Sample a state $\mathbf{x}^{\text{rand}} \in \mathcal{W}_{\text{free}}$ “randomly”;
- 2) Determine a new state \mathbf{x}^{new} to be added to the tree based on some distance criteria on \mathbf{x}^{rand} and \mathcal{T} ;
- 3) Determine set $\mathcal{Q}^{\text{near}} \subseteq \mathcal{V}_{\mathcal{T}}$ based on some distance criteria from which the tree will be extended to \mathbf{x}^{new} ;
- 4) For each state $q_P^{\text{near}} = (\mathbf{x}^{\text{near}}, q_B^{\text{near}}) \in \mathcal{Q}_P^{\text{near}}$, we consider a potential edge from q_P to $q_P^{\text{new}} = (\mathbf{x}^{\text{new}}, q_B)$ if $(\mathbf{x}^{\text{near}}, \mathbf{x}^{\text{new}}) \in \rightarrow_T$ and $(q_B^{\text{near}}, L(\mathbf{x}^{\text{near}}), q_B^{\text{new}}) \in \rightarrow_B$;
- 5) Add q_P^{new} to tree \mathcal{T} from state $q_P \in \mathcal{Q}_P^{\text{near}}$ to minimize $\text{Cost}(q_P^{\text{new}}) = \text{Cost}(q_P^{\text{near}}) + C(q_P^{\text{near}}, q_P^{\text{new}})$;
- 6) After adding q_P^{new} , the tree edges and costs are reconfigured so that each vertex is reached by a path with minimum cost from the root.

B. Sampling Strategies

In the construction of random trees, it remains to specify how state \mathbf{x}^{rand} is sampled “randomly”. In [21], the authors proposed two different sampling strategies for \mathbf{x}^{rand} :

- *Uniform Sampling*: \mathbf{x}^{rand} is selected according to a uniform distribution on $\mathcal{W}_{\text{free}}$ (in fact, any distribution as long as all states has non-zero probability);
- *Biased Sampling*: \mathbf{x}^{rand} is selected according to a biased distribution on $\mathcal{W}_{\text{free}}$ such that one has more chance to move towards an accepting state in B .

Since our NN-guided approach is related to the biased sampling, we briefly review this method. First, one needs to pre-process the NBA B so that infeasible transitions are removed. Then based on the simplified NBA, one defines a distance function $\rho : \mathcal{Q}_B \times \mathcal{Q}_B \rightarrow \mathbb{N}$ to capture the length of the shortest path between each pair of states. Then the selection of \mathbf{x}^{rand} is determined as follows:

- S-1 Select a *feasible accepting state* $q_B^{F, \text{feas}} \in \mathcal{Q}_B^F$ such that $\rho(q_B^0, q_B^{F, \text{feas}}) \neq \infty$ and $\rho(q_B^{F, \text{feas}}, q_B^{F, \text{feas}}) \neq \infty$;
- S-2 Define the set of vertices $\mathcal{D}_{\text{min}} \subseteq \mathcal{V}_{\mathcal{T}}$ that are closest to $q_B^{F, \text{feas}}$ according the distance function ρ . Then select a vertex $q_P^{\text{closest}} = (\mathbf{x}^{\text{closest}}, q_B^{\text{closest}}) \in \mathcal{V}_{\mathcal{T}}$ from the tree according to a specific distribution [21] such that states in \mathcal{D}_{min} has more chance to be selected;
- S-3 Select two successive states $q_B^{\text{succ}, 1}, q_B^{\text{succ}, 2} \in \mathcal{Q}_B$ such that $q_B^{\text{closest}} \xrightarrow{L(\mathbf{x}^{\text{closest}})} q_B^{\text{succ}, 1} \rightarrow q_B^{\text{succ}, 2}$, and $\rho(q_B^{\text{succ}, 2}, q_B^{\text{feas}}) \leq \rho(q_B^{\text{succ}, 1}, q_B^{\text{feas}}) \leq \rho(q_B^{\text{closest}}, q_B^{\text{feas}})$;
- S-4 Select a state $\mathbf{x}^{\mathcal{L}}$ such that $q_B^{\text{succ}, 1} \xrightarrow{L(\mathbf{x}^{\mathcal{L}})} q_B^{\text{succ}, 2}$ is feasible;

- S-5 Compute the shortest path from $\mathbf{x}^{\mathcal{L}}$ and $\mathbf{x}^{\text{closest}}$ within $\mathcal{W}_{\text{free}}$. Pick the second point in the shortest path, denoted by $\mathbf{x}^{\text{target}}$ as a heuristic direction for sampling;
- S-6 Finally, select \mathbf{x}^{rand} according to a distribution that has more chance to sample towards the direction of $\mathbf{x}^{\text{target}}$.

IV. NEURAL NETWORK GUIDED SAMPLING

Although biased sampling provides a more efficient approach compared with the uniform sampling, it solely relies on the distance information in Büchi automaton. For example, when $q_B^{F, \text{feas}}$ as well as intermediate states $q_B^{\text{succ}, 1}$ and $q_B^{\text{succ}, 2}$ are selected, one only consider whether or not the task progress can be pushed forward in the NBA. However, the information from the workspace, i.e., whether the task progress is feasible physically, is neglected.

To further improve the sample efficiency for the random tree construction, one essentially needs to *jointly* consider the information workspace and the NBA in order to have more chance to sample states that can move towards accepting states *by paths that are feasible in the workspace*. However, obtaining such a good heuristic is very challenging since the workspace is continuous. In this section, we present a new neural network guided (NN-guided) sampling strategy, where a “sample net” is used that effectively fuses the information of the workspace and the NBA. All implementation details as well as source codes are available at <https://github.com/LRJ-2000/NNgTL>

A. Overview of NN-Guided Approach

1) *Purposes of Network Networks*: The main architecture of our sample net is shown in Figure 1 consisting of two sub-networks: the Path Prediction Network (PathNet) and the State Prediction Network (StateNet). The inputs of the PathNet and the StateNet are the same, which are both the workspace map as well as the NBA for the LTL task. However, their purposes and outputs are different. Specifically,

- The output of the StateNet is a vector \mathbf{p} with a length of $|\mathcal{Q}_B|$. In this vector, each entry p_i denotes the probability that state i is involved in the optimal path. The prediction vector \mathbf{p} is employed to guide the choices of $q_B^{F, \text{feas}}, q_B^{\text{succ}, 1}, q_B^{\text{succ}, 2} \in \mathcal{Q}_B$ in the biased sampling.
- The output of the PathNet is a 200×200 matrix \mathcal{P} such that the value of each entry represents the likelihood that the entry is on the optimal path. Therefore, this weight matrix \mathcal{P} is used as a more reasonable metric for sampling \mathbf{x}^{rand} by considering the workspace information with preforming shortest path search.

2) *NN-Guided Sampling Strategy*: Now, let us discuss how the proposed neural networks are used to guide the sampling process. It still follows the basic idea of biased sampling method as detailed in steps S1-S6. However, we further leverage the outputs of the two sub-networks to improve and simplify the sampling process. Specifically, at each instant, suppose that the output of the StateNet and the PathNet are \mathbf{p} and \mathcal{P} . We set $\alpha \in (0, 1)$ as a parameter that specifies the probability of using the predicted information of StateNet to guide the sampling. Then, to

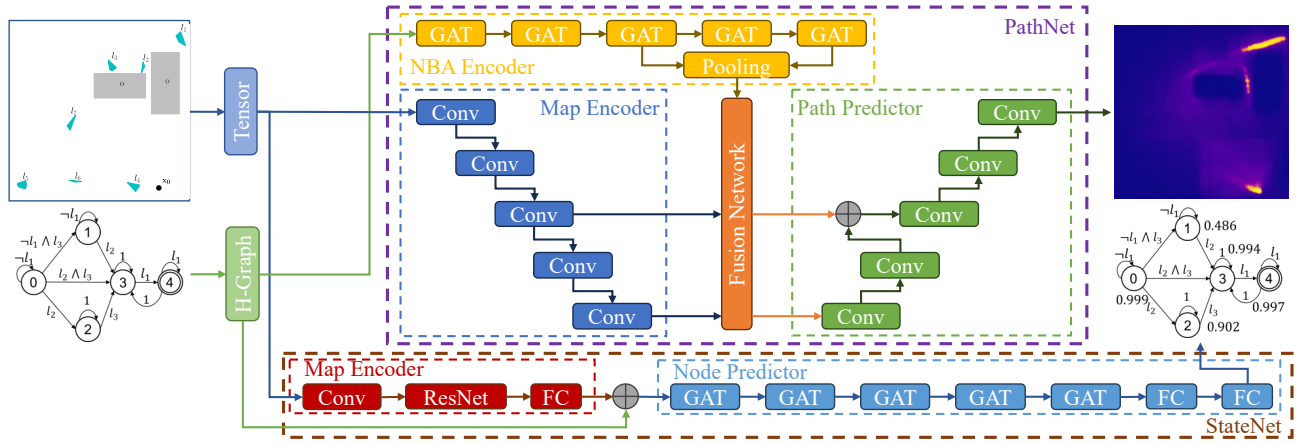


Fig. 1: Overview of the sampling network.

determine the sample point \mathbf{x}^{rand} , our approach makes the following changes C-1 and C-2 to steps S-1 to S-6.

C-1: When selecting $q_B^{F, \text{feas}}$, q_B^{closest} , $q_B^{\text{succ}, 1}$ as in steps S1-S3 in the biased sampling, there is a $1 - \alpha$ probability that we still follow exactly the same strategy in S-1 to S-3. However, there is an α probability that all these states selected according to the probability vector \mathbf{p} . In other words, we have α probability to activate the prediction result of the StateNet.

C-2: After obtaining state $\mathbf{x}^{\mathcal{L}}$ in step S-4, to sample state \mathbf{x}^{rand} , we simplify steps S-5 and S-6 as a single step. Particularly, instead of first computing the shortest path and then using the second point in the path to generate a sample distribution, here we directly use the prediction result \mathcal{P} by the PathNet to sample \mathbf{x}^{rand} . More specifically, let (x_1, y_1) and (x_2, y_2) be the coordinates of $\mathbf{x}^{\mathcal{L}}$ and $\mathbf{x}^{\text{closest}}$, respectively. We consider the rectangle region $\text{REC} = [x_1 : x_2, y_1 : y_2]$. We define a discrete distribution over all grids in REC according to the normalized value of their weights in \mathcal{P} , i.e., grids with larger value has more chance to be selected. Then \mathbf{x}^{rand} is sampled randomly from the rectangle region according to this distribution.

Here, we discuss the main features of the proposed NN-guided sampling strategy. First, our NN-guided approach subsumes all properties, such as probabilistic completeness and asymptotic optimality, of the TL-RRT* algorithm in [21] since we still follow the main structure of TL-RRT* and at each step, our algorithm has a non-zero probability to switch to the original sampling strategy. However, compared with the biased sampling strategy adopted in [21], our NN-guided sampling strategy further jointly considers both the workspace information and the NBA, which provides a better heuristic for “good” samples. Furthermore, since our strategy uses the predicted distribution directly without involving the shortest path search for each step, its online execution is also much faster than the biased sampling strategy.

B. Inputs Encodings

Recall that, for both the PathNet and the StateNet, their inputs are the workspace map and the NBA for the LTL formula. To leverage neural networks for processing the

workspace with continuous space and NBA with graph structure, appropriate encoding techniques are needed.

1) *Encoding for Workspace*: First, we consider the continuous workspace as a 200×200 pixel image or a grid map. Each grid in the image corresponds to a specific point in the workspace, which is labeled, obstacle or free. Then the image is transformed into a tensor of dimensions $(m+1) \times 200 \times 200$ with m be the number of different labels, i.e., each grid in the workspace is encoded as a vector $\mathbf{a} = [a_0, a_1, \dots, a_m]$. Specifically, the first entry a_0 represents the grid’s status, where -1 for the initial location, 0 for free space, and 1 for an obstacle. For $i = 1, \dots, m$, we have $a_i = 1$ if it belongs to \mathcal{R}_i , and $a_i = 0$ otherwise.

2) *Encoding for Büchi Automata*: First, we convert NBA to a directed graph, where nodes and edges correspond to states and transitions, respectively, and have their features. Specifically, the feature for a node is a vector $\mathbf{v} = [v_1, v_2, v_3]$, where $v_1 \in \{0, 1\}$ represents whether or not it is an initial state, $v_2 \in \{0, 1\}$ represents whether or not it is an feasible accepting state, and v_3 is the (normalized) distance to the closest feasible accepting state. For each edge, the feature is a vector $\mathbf{e} = [e_1, e_2, \dots, e_m] \in \{-1, 0, 1\}^m$ specifying the atomic propositions that need to be true for the underlying transition in the NBA. Since later on we use graph neural networks (GNN) to process features of the NBA, we further transform the directed graph into a heterogeneous graph. This transformation involves adding a new node within each original edge so that the features of the edges are inherited by the nodes added. Additionally, to augment feature aggregation and spread, a self-loop is added to each node, and a pooling node is added so that each node has a directed edge leading to this pooling node.

C. Implementation Details of Neural Networks

1) *PathNet*: It has the following four building blocks.

Map Encoder: The purpose is to extract features from the grid map by five convolutional blocks. Each block houses a 4×4 convolutional layer with a stride of 2 and padding of 1, complemented by a batch-norm layer, a dropout layer with probability 0.5, and a LeakyReLU activation with a negative slope of 0.2. Note that the sizes of the map evolves from

$m \times 200 \times 200$ to $1024 \times 6 \times 6$ when passing through these blocks.

NBA Encoder: The purpose is to extract global features from the NBA. Comprising five layers, the encoder utilizes Graph Attention (GAT) convolutions to address distinct edge types in the heterogeneous graph [32]. The node features are proceeded by a dropout layer and a ReLU activation after convolutions. Finally, global mean pooling operations are used to accumulate feature representation of the entire graph.

Fusion Network: The purpose is to amalgamate features from both the workspace Map Encoder and the NBA Encoder. This is done by the following two steps. First, NBA features are transformed via a linear layer to attain compatibility with map features. Then vector concatenation are used to fuse these harmonized features.

Path Predictor: The purpose is to output the weight matrix $\mathcal{P} \in \mathbb{R}^{200 \times 200}$ as the prediction for optimal paths. To this end, we use five up-convolution blocks to upscale and to refine the amalgamated features. Each block is structured with a transposed convolutional layer, a batch-norm layer, a dropout layer (with probability 0.5), and a ReLU activation. Drawing from the “U-Net” architecture [33], our design integrates skip connections, merging features from both the third convolutional modules of the Map and NBA Encoders. This approach capitalizes on the synergy of both encoders, enhancing the merged feature representation. Subsequent to this fusion, features are relayed to the Path Predictor via concatenation. Leveraging these connections ensures the retention of intricate spatial details alongside the depth of hierarchical features. Finally, we use an 1×1 convolution to reshape the features to $1 \times 200 \times 200$ dimensions, and a sigmoid activation is used to refine the output.

2) *StateNet*: This net is essentially a classifier for the NBA states. It has a simpler structure consisting of the following two components.

Map Encoder: The main purpose is to encode the map into a 256-length feature vector, which will be served as a precursor to the Node Predictor. Specifically, the map is processed by an 1×1 convolution, transitioning the 8-channel input to 3 channels, thus aligning with conventional image processing frameworks. The features then engage with the pre-trained “ResNet-50” model [34]. The terminal fully connected layer of the ResNet is omitted, and is replaced by our bespoke fully connected layer, rendering the output as a 256-length feature vector.

Node Predictor, tailored for node classification, starts with a GAT convolution layer, stretching the graph’s node features to a 256-length dimension, in line with the map features. Each node’s features are subsequently concatenated with the Map Encoder-generated map feature vector, amalgamating spatial and structural data at every node. A sequence of information dissemination follows via five sequential GAT modules, each housing a GAT convolution, a ReLU activation for non-linearity, and a dropout (with probability 0.5) for regularization. The classification culminates with two sequential fully connected layers and a Softmax layer,

analyzing the consolidated node features to yield the classification probability.

D. Training Neural Networks

1) *Data Set Preparations*: Initially, we randomly generate 15400 pairs of workspace and LTL formulae. Specifically, in each 200×200 grid map workspace, we randomly place obstacles as well as seven distinct labeled regions. The initial location of the robot is also randomized. Note that grid map are only used for the purpose of map generation, it is still mapped to and used as continuous workspace. To order to obtain an expert path for each pair of workspace and LTL task, we use the existing biased-sampling approach with 10,000 iterations. The obtained expert path is then encoded into a 200×200 binary matrix. Specifically, we mark those grids crossed by the expert path, as well as their immediate neighbors, as 1. Then we label NBA states visited by the expert path as 1. Finally, through data augmentations, this data set is expanded to 107800 cases.

2) *Training Procedures*: The training process started with training the StateNet first. Specifically, we first fixed the parameters of the ResNet portion and only update the parameters of the other parts. When the loss became stable after 100 epochs, we unfroze the parameters of the ResNet and continued training for another 20 epochs. The trainings were performed by the Adam optimizer with an initial learning rate of 0.001 and a batch size of 128. We used cross entropy loss as the loss function. After training the StateNet, we trained the PathNet. Specifically, we first initialized the GAT layers in the NBA Encoder using the parameters from the GAT layers in StateNet. After training for 150 epochs, the loss stabilized. We still used the Adam optimizer with an initial learning rate of 0.0001 and a batch size of 128. We employed the binary cross entropy loss as the loss function.

V. SIMULATIONS & NUMERICAL EXPERIMENTS

In this section, we provide simulation results for the proposed method. First, a case study is provided to illustrate our approach. Then we perform a set of numerical experiments to evaluate the efficiency of our approach compared with existing sampling strategies. All algorithms were implemented using Python on a Windows 11 computer with an Intel Core i7-13700K 5.40GHz processor.

A. Case Study

We consider a robot moving in a workspace shown in Figure 2 with initial position of the robot, obstacles, and labeled regions as depicted in the figure. We consider the following LTL task for the robot

$$\phi = \square \Diamond l_1 \wedge \neg l_1 \mathcal{U} l_2 \wedge \Diamond l_3,$$

i.e., the robot needs to (i) visit l_2 at least once without visiting l_1 ; (ii) visit l_3 at least once; (iii) visit l_1 infinitely often. The NBA of ϕ is shown in Figure 3b, where state 4 is the unique feasible accepting state. The optimal plan searched by our algorithm is shown as the red lines in Figure 2. The robot first goes to l_2 , then goes to l_3 , and finally stays at l_1 . That is, only the prefix part contributes to the overall cost.

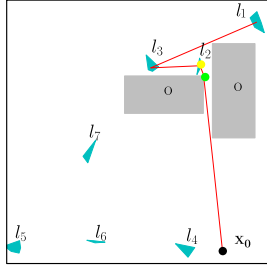
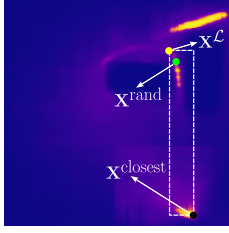
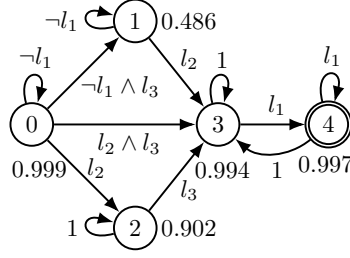


Fig. 2: Workspace of the robot, where the gray regions denote obstacles, the green regions denote labeled regions, and the red lines denote the optimal path synthesized.



(a) Output of PathNet.



(b) Output of StateNet.

Fig. 3: Predication results of the neural networks.

To better illustrate our NN-guided sampling strategy, we explain how the random tree \mathcal{T} for the prefix path is expanded initially from the root $q_P^0 = (x_0, q_B^{\text{init}})$. Since the NBA only has one feasible accepting state, the choice for $q_B^{F, \text{feas}}$ is unique, and we have $q_P^{\text{closest}} = q_P^0$ since the tree only has a root so far. Then the StateNet predicts $\mathbf{p} = [0.999, 0.486, 0.902, 0.994, 0.997]$. Since $0 \xrightarrow{L(x_0)}_B 0$, we have $q_B^{\text{succ}, 1} = 0$. From the feasible successor states of $q_B^{\text{succ}, 1}$, we choose $q_B^{\text{succ}, 2} = 2$ since it has a higher probability (0.902) than state 1 (0.486). Then according to the transition condition $0 \xrightarrow{l_2}_B 2$ in the NBA, we select a point in the labeled region l_2 as x^L , which is shown as the yellow point in Figure 3. Then we construct the rectangle region between x^L and x^{closest} , and state x^{rand} is sampled randomly according to the distribution determined by the PathNet prediction \mathcal{P} as shown in Figure 3a.

Note that, in this example, there are actually two feasible prefix paths: $l_2 \rightarrow l_3 \rightarrow l_1$ and $l_3 \rightarrow l_2 \rightarrow l_1$. Clearly, the former is optimal with less cost. This information is captured by the prediction result of the StateNet, which prefers state $2 \in \mathcal{Q}_B$ for the optimal path. Furthermore, the prediction result of the PathNet can effectively avoid obstacles and has more chance to sample near the optimal path. Therefore, these prediction results by the neural networks guide our search process to converge to a desired solution more quickly.

B. Comparison with Existing Methods

1) *Experiment Settings:* We conduct a set of experiments to illustrate the efficiency of our NN-guided sampling strategy compared with the existing uniform and biased sampling strategies. Specifically, independent from the training set, we generate another 240 pairs of workspace map and LTL task. For each instance, we run our method with $\alpha = 0.8$ as

TABLE I: Experiment Results

(a) Statistic Results for Simple Tasks

Method	$T(s)$	n	len	m
Uniform	54.2475	1908.13	0.78789	3167.63
Biased	1.03305	101.000	0.80842	200.307
NN-Guided	0.09518	22.8941	0.90546	48.1395

(b) Statistic Results for Complex Tasks

Method	$T(s)$	n	len	m
Uniform	1401.78	7401.27	-	21501.2
Biased	15.5639	321.553	1.08553	704.082
NN-Guided	2.19490	44.8549	1.12555	182.743

well as the two existing method to find a feasible plan. Note that, since all these RRT-based approaches are probabilistic optimal, we focus on comparing *how fast* these strategies can find the *first feasible solution*, and the performance of the first feasible solution in terms of its length. Formally, we consider the following metrics when the first feasible solution is found: the execution time T taken, the number of iterations n required, the number of nodes m in the random tree, and the length of the first feasible solution len .

2) *Statistic Results:* The numerical experiments results are shown in Table I. Specifically, based on the execution time T of the uniform sampling approach, we divide the tasks into simple tasks with $T \leq 180s$ and complex tasks with $T > 180s$. Then Tables Ia and Ib show the average value of each metric for each algorithm within these two task categories, respectively. Note that many complex tasks, the uniform sampling strategy fails to find a feasible path within 2000s. For such cases, we terminate the search by only recording T, n and m without consider the length len .

The statistic results show that, for both simple and complex tasks, our NN-guided sampling strategy can significantly enhance the efficiency of the RRT-based algorithm. In particular, the timed required to obtain a feasible solution by our NN-guided sampling strategy is less than 15% of that of the biased sampling strategy, which is already much more efficient than uniform sampling strategy. Furthermore, the length of the first feasible solution found by our strategy is similar to those of the other two strategies. We would like the remark that the metric of len is not essential compared with other metrics since this is just the performance of the first feasible solution. The entire algorithm will converge to the optimal solution when the number of iterations increases.

VI. CONCLUSION

In this paper, building upon the current state-of-the-art sampling-based LTL planning algorithms, we propose a novel sampling strategy based on multi-modal neural networks to guide the sampling process. Our approach, on the one hand, leverages the feature extraction power of neural networks in an end-to-end fashion, and on the other hand, still enjoys all good properties of the sampling-based methods such as probabilistic optimality/completeness. Experimental results show that our proposed sampling strategy can significantly enhance the planning efficiency of the algorithm. In future research, we will further improve the feature fusion methods for multi-robot planning problems.

REFERENCES

- [1] M. Guo and M. M. Zavlanos, "Probabilistic motion planning under temporal tasks and soft constraints," *IEEE Transactions on Automatic Control*, vol. 63, no. 12, pp. 4051–4066, 2018.
- [2] M. Cai, H. Peng, Z. Li, and Z. Kan, "Learning-based probabilistic ltl motion planning with environment and motion uncertainties," *IEEE Transactions on Automatic Control*, vol. 66, no. 5, pp. 2386–2392, 2022.
- [3] X. Luo and M. M. Zavlanos, "Temporal logic task allocation in heterogeneous multirobot systems," *IEEE Transactions on Robotics*, vol. 38, no. 6, pp. 3602–3621, 2022.
- [4] X. Yu, X. Yin, S. Li, and Z. Li, "Security-preserving multi-agent coordination for complex temporal logic tasks," *Control Engineering Practice*, vol. 123, p. 105130, 2022.
- [5] S. Li, M. Wei, S. Li, and X. Yin, "Temporal logic task planning for autonomous systems with active acquisition of information," *IEEE Transactions on Intelligent Vehicles*, vol. 9, pp. 1436–1449, 2024.
- [6] S. L. Smith, J. Tůmová, C. Belta, and D. Rus, "Optimal path planning for surveillance with temporal-logic constraints," *The International Journal of Robotics Research*, vol. 30, no. 14, pp. 1695–1708, 2011.
- [7] E. Plaku and S. Karaman, "Motion planning with temporal-logic specifications: Progress and challenges," *AI Communications*, vol. 29, no. 1, pp. 151–162, 2016.
- [8] Y. Yang, J. Zhao, X. Yin, and S. Li, "Formal synthesis of warehouse robotic systems with temporal logic specifications," in *IEEE Conference on Robotics, Automation and Mechatronics*, 2019, pp. 469–474.
- [9] G. Scher and H. Kress-Gazit, "Warehouse automation in a day: from model to implementation with provable guarantees," in *IEEE International Conference on Automation Science and Engineering*, 2020, pp. 280–287.
- [10] M. Guo and D. V. Dimarogonas, "Multi-agent plan reconfiguration under local ltl specifications," *The International Journal of Robotics Research*, vol. 34, no. 2, pp. 218–235, 2015.
- [11] M. Kloetzer and C. Belta, "Automatic deployment of distributed teams of robots from temporal logic motion specifications," *IEEE Transactions on Robotics*, vol. 26, no. 1, pp. 48–61, 2009.
- [12] W. Shi, Z. He, W. Tang, W. Liu, and Z. Ma, "Path planning of multi-robot systems with Boolean specifications based on simulated annealing," *IEEE Robotics and Automation Letters*, vol. 7, no. 3, pp. 6091–6098, 2022.
- [13] A. Ulusoy, S. L. Smith, and C. Belta, "Optimal multi-robot path planning with ltl constraints: guaranteeing correctness through synchronization," in *Distributed Autonomous Robotic Systems: The 11th International Symposium*, 2014, pp. 337–351.
- [14] J. Zhao, K. Zhu, S. Li, and X. Yin, "To explore or not to explore: Regret-based ltl planning in partially-known environments," *IFAC-PapersOnLine*, vol. 56, no. 2, pp. 11 337–11 343, 2023.
- [15] P. Lv, G. Luo, Z. Ma, S. Li, and X. Yin, "Optimal multi-robot path planning for cyclic tasks using petri nets," *Control Engineering Practice*, vol. 138, p. 105600, 2023.
- [16] S. Karaman and E. Frazzoli, "Sampling-based algorithms for optimal motion planning," *The International Journal of Robotics Research*, vol. 30, no. 7, pp. 846–894, 2011.
- [17] —, "Sampling-based algorithms for optimal motion planning with deterministic μ -calculus specifications," in *American Control Conference*, 2012, pp. 735–742.
- [18] C. I. Vasile and C. Belta, "Sampling-based temporal logic path planning," in *International Conference on Intelligent Robots and Systems*, 2013, pp. 4817–4822.
- [19] Y. Kantaros and M. M. Zavlanos, "Sampling-based optimal control synthesis for multirobot systems under global temporal tasks," *IEEE Transactions on Automatic Control*, vol. 64, no. 5, pp. 1916–1931, 2018.
- [20] —, "Stylus*: A temporal logic optimal control synthesis algorithm for large-scale multi-robot systems," *The International Journal of Robotics Research*, vol. 39, no. 7, pp. 812–836, 2020.
- [21] X. Luo, Y. Kantaros, and M. M. Zavlanos, "An abstraction-free method for multirobot temporal logic optimal control synthesis," *IEEE Transactions on Robotics*, vol. 37, no. 5, pp. 1487–1507, 2021.
- [22] B. Ichter, J. Harrison, and M. Pavone, "Learning sampling distributions for robot motion planning," in *IEEE International Conference on Robotics and Automation*, 2018, pp. 7087–7094.
- [23] J. Wang, W. Chi, C. Li, C. Wang, and M. Q.-H. Meng, "Neural rrt*: Learning-based optimal path planning," *IEEE Transactions on Automation Science and Engineering*, vol. 17, no. 4, pp. 1748–1758, 2020.
- [24] N. Ma, J. Wang, J. Liu, and M. Q.-H. Meng, "Conditional generative adversarial networks for optimal path planning," *IEEE Transactions on Cognitive and Developmental Systems*, vol. 14, no. 2, pp. 662–671, 2021.
- [25] A. H. Qureshi, Y. Miao, A. Simeonov, and M. C. Yip, "Motion planning networks: Bridging the gap between learning-based and classical motion planners," *IEEE Transactions on Robotics*, vol. 37, no. 1, pp. 48–66, 2020.
- [26] W. Liu, N. Mehdipour, and C. Belta, "Recurrent neural network controllers for signal temporal logic specifications subject to safety constraints," *IEEE Control Systems Letters*, vol. 6, pp. 91–96, 2021.
- [27] W. Hashimoto, K. Hashimoto, and S. Takai, "Stl2vec: Signal temporal logic embeddings for control synthesis with recurrent neural networks," *IEEE Robotics and Automation Letters*, vol. 7, no. 2, pp. 5246–5253, 2022.
- [28] M. Cai, M. Hasanbeig, S. Xiao, A. Abate, and Z. Kan, "Modular deep reinforcement learning for continuous motion planning with temporal logic," *IEEE Robotics and Automation Letters*, vol. 6, no. 4, pp. 7973–7980, 2021.
- [29] J. J. Alloor, J. Patrikar, P. Kapoor, J. Oh, and S. Scherer, "Follow the rules: Online signal temporal logic tree search for guided imitation learning in stochastic domains," in *IEEE International Conference on Robotics and Automation*, 2023, pp. 1320–1326.
- [30] M. Kloetzer and C. Belta, "A fully automated framework for control of linear systems from temporal logic specifications," *IEEE Transactions on Automatic Control*, vol. 53, no. 1, pp. 287–297, 2008.
- [31] C. Baier and J.-P. Katoen, *Principles of model checking*. MIT press, 2008.
- [32] P. Velickovic, G. Cucurull, A. Casanova, A. Romero, P. Lio, Y. Bengio *et al.*, "Graph attention networks," *Stat*, vol. 1050, no. 20, pp. 10–48 550, 2017.
- [33] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *Medical Image Computing and Computer-Assisted Intervention*, 2015.
- [34] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 770–778.