Sampling-based Motion Planning with Symbolic, Geometric, and Differential Constraints*

Erion Plaku, Gregory Hager
Department of Computer Science
Johns Hopkins University, Baltimore, MD 21218
{erion,hager}@cs.jhu.edu

Abstract

This paper presents a multi-layered approach (SamplSGD) to automatically compute collision-free and dynamically-feasible trajectories that satisfy high-level specifications given in a planning-domain definition language. A crucial aspect of SamplSGD is an interplay between sampling-based motion planning and symbolic action planning. SamplSGD leverages from sampling-based motion planning the underlying idea of searching for a solution trajectory by selectively sampling and exploring the continuous space of collision-free and dynamically-feasible motions. Drawing from AI, SamplSGD uses symbolic action planning to identify regions of the continuous space that sampling-based motion planning can further explore to significantly advance the search. The planning layers in SamplSGD interact with each-other through estimates on the utility of each action, which are computed based on information gathered during the search. Simulation experiments on a challenging object-manipulation task provide promising initial validation.

1. Introduction

Research in robotics has focused since its inception towards increasing the ability of robots to plan and act on their own in order to complete assigned high-level tasks. Toward this goal, this paper studies the following problem:

> Given a high-level specification, automatically plan the sequence of motions the robot needs to execute so that the resulting trajectory is dynamically feasible, avoids collisions, and satisfies the high-level specification.

There are two crucial aspects to this fundamental planning problem: (i) planning in the space of possible high-level actions and (ii) planning in the space of possible motions.

Action planning, which assumes a discrete world and discrete actions, has been extensively studied in AI and logic. Throughout the years, significant progress has been made in addressing increasingly complex discrete planning problems. In fact, current methods based on symbolic reasoning have made it possible to specify high-level goals using sophisticated planning-domain languages, such as STRIPS (Fikes and Nilsson 1971), ADL (Pednault 1994), PDDL (Ghallab et al. 1998), HAL (Marthi, Russell, and Wolfe 2007), and efficiently plan the sequence of discrete actions that accomplishes the specified goals even in discrete spaces with billions of states (see summaries in (Ghallab, Nau, and Traverso 2004; Baier and Katoen 2008)).

In contrast, motion planning assumes a continuous world and continuous motions. The motivation comes from navigation, exploration, search-and-rescue missions, and other applications where it is essential to compute trajectories that can be followed by the robot in the physical world. As a result, the planned motions need to not only avoid collisions with obstacles but also satisfy differential constraints imposed by the underlying robot dynamics. Due to the increased complexity, motion planning has generally been limited to simpler goal specifications, such as reachability, where the objective is to compute a collision-free and dynamically-feasible trajectory from an initial to a goal state (Choset et al. 2005; LaValle 2006).

Researchers have generally considered action planning and motion planning separately. As a result, the problem of planning motions that satisfy high-level specifications is typically approached by first using action planning to compute a sequence of discrete actions that satisfies the goal specification. In a second step, motion planning based on controllers is used to consecutively follow the discrete actions in the continuous world (Arkin 1990; Payton, Rosenblatt, and Keirsey 1990; Saffiotti, Konolige, and Ruspini 1995; Belta et al. 2007; Fainekos et al. 2007).

There are, however, several limitations to these decoupled approaches. Since discrete actions could have different meanings (e.g., "MoveTo," "PickUp," "PlaceOnTop"), numerous controllers would have to be designed in order to handle the potential diversity of the available discrete actions and robot dynamics. Moreover, controllers are not always available. In many cases, collision-avoidance requirements and differential constraints imposed by dynamics make it difficult or impossible to design a controller that can follow a discrete action in the continuous world. As a result, decoupled approaches have had limited success.

To efficiently plan collision-free and dynamically-feasible trajectories that satisfy high-level specifications, this paper treats it as a search problem over the discrete space of actions and the continuous space of motions. A multi-layered approach is presented, SamplSGD (*Sa*mpling-based *M*otion *Pl*anning with *S*ymbolic, *G*eometric, and *D*ifferential constraints), which combines action planning

*This work is supported by NSF IIS-0748338.

with sampling-based motion planning. SamplSGD leverages from sampling-based motion planning the underlying idea of searching for a solution trajectory by selectively sampling and exploring the continuous space of motions. Sampling-based motion planners are widely applicable and have had significant success in solving challenging reachability motion-planning problems in high-dimensional continuous spaces (Kavraki et al. 1996; Amato et al. 1998; LaValle and Kuffner 2001; Hsu et al. 2002; Sánchez and Latombe 2002; Ladd and Kavraki 2005; Plaku, Kavraki, and Vardi 2007; 2008b; Choset et al. 2005; LaValle 2006). To handle both collision-avoidance requirements and differential constraints imposed by dynamics, SamplSGD uses a tree-based exploration of the continuous space. The tree is rooted at the initial state and is incrementally extended with trajectories obtained by applying input controls to the states in the tree and propagating the dynamics forward in time. The success and computational efficiency of the tree-based exploration depends on the ability of SamplSGD to effectively guide the tree-based exploration toward the goal, and, as a result, add as quickly as possible a vertex v such that the trajectory from the root to v satisfies the high-level specification. Drawing from AI, SamplSGD uses symbolic action planning to guide the tree-based exploration by identifying and selecting discrete actions and regions of the continuous space that sampling-based motion planning can further explore to significantly advance the search for a solution trajectory. The planning layers in SamplSGD interact with eachother through estimates on the utility of discrete states and actions, which are computed based on information gathered during the tree-based exploration. Thus, the symbolic action planning guides the sampling-based motion planning, while the latter feeds back information in the form of utility estimates to improve the guide in the next iteration. This interplay between symbolic action planning and sampling-based motion planning through the utility estimates allows SamplSGD to make proper use of the computational time. SamplSGD becomes increasingly successful in identifying regions whose further exploration can significantly advance the search while avoiding spending valuable computational time exploring regions that do not advance the search. Simulation experiments provide promising initial validation.

SamplSGD is motivated by earlier work on manipulation planning (Alami, Laumond, and Siméon 1995; Nielsen and Kavraki 2000; Gravot, Cambon, and Alami 2003; Cambon, Gravot, and Alami 2004; Cambon, Alami, and Gravot 2009) and hybrid systems (Plaku, Kavraki, and Vardi 2007; 2008b; 2008a; 2009). The work in (Nielsen and Kavraki 2000) used discrete search over the manipulation graph to guide a PRM (Probabilistic RoadMap (Kavraki et al. 1996)) sampling-based planner in the computation of transfer and transit paths. Later work by (Gravot, Cambon, and Alami 2003; Cambon, Gravot, and Alami 2004; Cambon, Alami, and Gravot 2009) led to the aSyMov planner, which extended the idea even further by combining PRM with symbolic action planning, making it possible to specify high-level goals in planning-domain definition languages.

A limitation of these approaches that rely on roadmaps is that they cannot take into account differential constraints imposed by dynamics. To construct a roadmap, each edge (a, b) of the roadmap requires connecting the state a to b via a trajectory that satisfies differential constraints imposed by dynamics. Exact solutions to this steering problem are available only in limited cases, while numerical solutions impose significant computational cost (Keller 1992), rendering roadmap construction impractical.

In contrast, SamplSGD uses a tree-based exploration of the state space, which does not require any steering, but only the ability to propagate dynamics forward in time. Forward propagation is readily achieved through numerical integration, making it possible for SamplSGD to generate not only collision-free but also dynamically-feasible trajectories that satisfy the high-level specification. Moreover, an essential component of aSyMov is a computationally-intensive backtracking procedure that checks for collisions to ensure that a candidate action is grounded in a geometric context that has at least a collision-free path from the initial state. SamplSGD takes a different approach, which avoids backtracking, by validating (checking for collisions) each trajectory before adding it to the tree.

SamplSGD builds upon earlier work (Plaku, Kavraki, and Vardi 2007; 2008b; 2008a; 2009), which showed how to effectively combine sampling-based motion planning with discrete search to compute collision-free and dynamically-feasible trajectories that satisfy high-level specifications given by linear temporal logic. A limitation of the work in (Plaku, Kavraki, and Vardi 2007; 2008b; 2008a; 2009), is that it relies on an explicit representation of the discrete space and the possible transitions between the discrete states. To address this limitation, SamplSGD integrates sampling-based motion planning with symbolic action planning, which can handle complex discrete planning problems. The integration of sampling-based motion planning with symbolic action planning also makes it possible for SamplSGD to handle high-level goal specifications given by planning-domain languages.

2. Preliminaries

Discrete Specifications by Planning-Domain Definition Languages

Drawing from research in AI, this paper uses planning-domain definition languages, such as STRIPS, to allow for sophisticated high-level planning specifications. Details can be found in standard AI books (Russell and Norvig 2002; Ghallab, Nau, and Traverso 2004). For completeness, a summary follows. A discrete model is a tuple $\mathcal{M} = (\mathcal{O}, \mathcal{P}, \mathcal{Q}, q_{\text{init}}, \phi_{\text{goal}}, \mathcal{A})$, where

- $\mathcal{O}$ denotes the set of objects (or atoms).
- $\mathcal{P}$ denotes the set of predicates, which express relations among objects in $\mathcal{O}$.
- $\mathcal{Q}$ denotes the discrete state space. A discrete state is a conjunction of all positive and grounded literals that currently hold in the world. A positive literal is of the form $P(t_1, \ldots, t_m)$, where $P \in \mathcal{P}$ is a predicate and each t_i is an object or an object variable. A positive literal is grounded if it does not contain any object variables.

- $q_{\text{init}} \in \mathcal{Q}$ denotes the initial discrete state of the world.
- ϕ_{goal} denotes the goal specification, which is given as a formula constructed by combining positive grounded literals with Boolean operators $\neg$, $\vee$, and $\wedge$.
- $\mathcal{A}$ denotes the set of action schemas. An action schema $A = (\text{vars}, \text{pre}, \text{post}) \in \mathcal{A}$ is defined in terms of object variables, a precondition that must hold before execution, and a postcondition that will hold after execution. A precondition is usually given as a conjunction of positive literals, while a postcondition is given as a conjunction of positive or negative literals. An action $a \in A$ is a specific instantiation of the variables in A. If the precondition is satisfied, the execution of a changes the current state according to the postcondition, i.e., adding positive literals and deleting negative literals. If the precondition is not satisfied, a has no effect.

A discrete solution consists of a finite sequence of actions $[a_i]_{i=1}^{n}$ that transforms the world from q_{init} to a discrete state that satisfies ϕ_{goal}.

Interpretation over Continuous Spaces

The physical world, which includes the robotic system, obstacles, and objects to be manipulated, is commonly modeled in a continuous setting. The continuous space of the world, denoted by $\mathcal{S}$, consists of a finite collection of continuous variables that can describe the world, e.g., placement of objects, joint values in a robot arm, vehicle velocity.

The continuous space $\mathcal{S}$ gives meaning to the predicates in the discrete specification. As an example, $\text{On}(\text{book}, \text{table})$ holds iff the book is actually on the table. Since a continuous state $s \in \mathcal{S}$ specifies the placement of the objects, one can determine whether or not the predicate holds at s. This interpretation of which predicates actually hold at a continuous state provides a mapping from the continuous space to the discrete space, denoted as a function $\text{map}_{\mathcal{S}\mapsto\mathcal{Q}} : \mathcal{S} \to \mathcal{Q}$.

Moreover, trajectories over $\mathcal{S}$ give meaning to the actions in the discrete specification. A trajectory over $\mathcal{S}$ is a continuous function $\zeta : [0, T] \to \mathcal{S}$, parametrized by time. As the continuous state changes according to ζ, the discrete state, obtained by the mapping $\text{map}_{\mathcal{S}\mapsto\mathcal{Q}} : \mathcal{S} \to \mathcal{Q}$, may also change. As a result, the trajectory ζ *follows* a discrete action a if (i) $\text{map}_{\mathcal{S}\mapsto\mathcal{Q}}(\zeta(0))$ satisfies a's precondition and (ii) $\text{map}_{\mathcal{S}\mapsto\mathcal{Q}}(T)$ satisfies a's postcondition.

The underlying dynamics are specified as a set of differential equations $f : \mathcal{S} \times \mathcal{U} \to \dot{\mathcal{S}}$, where $\mathcal{U}$ is a control space consisting of a finite set of input variables that can be applied to the system (e.g., a car can be controlled by setting the acceleration and the rotational velocity of the steering wheel). A dynamically-feasible trajectory $\zeta : [0, T] \to \mathcal{S}$ is obtained by computing a control function $\tilde{u} : [0, T] \to \mathcal{U}$ and propagating the dynamics forward in time through numerical integration from a given state $s \in \mathcal{S}$, i.e.,

$$\zeta(t) = s + \int_0^t f(\gamma(h), \tilde{u}(h))\, dh.$$

The dynamically-feasible trajectory $\zeta : [0, T] \to \mathcal{S}$ is considered collision free if each state along the trajectory avoids collisions with the obstacles.

3. SamplSGD

To effectively compute a collision-free and dynamically-feasible trajectory that starts at $s_{\text{init}} \in \mathcal{S}$ and satisfies the discrete goal specification ϕ_{goal}, SamplSGD conducts the search both in the continuous space $\mathcal{S}$ and in the discrete state and action spaces, $\mathcal{Q}$ and $\mathcal{A}$.

In the continuous space $\mathcal{S}$, SamplSGD maintains the search data structure as a tree $\mathcal{T} = (V, E)$. Each vertex $v \in \mathcal{T}[V]$ is associated with some continuous state $s \in \mathcal{S}$, written as $v.s$. An edge $(v', v'') \in \mathcal{T}[E]$ indicates that SamplSGD has computed a collision-free and dynamically-feasible trajectory from $v'.s$ to $v''.s$.

Initially, $\mathcal{T}[V]$ contains only one vertex, v_{init}, which is associated with the initial state $s_{\text{init}} \in \mathcal{S}$, and $\mathcal{T}[E]$ is empty. As SamplSGD explores $\mathcal{S}$, new vertices and new edges are added to $\mathcal{T}$. The procedure consists of selecting a vertex $v \in \mathcal{T}$ for expansion and then extending the tree from v by generating a collision-free and dynamically-feasible trajectory that starts at $v.s$. A common strategy is to apply some control $u \in \mathcal{U}$ to $v.s$ and simulate the dynamics forward in time until a collision occurs, a state-constraint is violated, or a maximum number of steps is exceeded (Choset et al. 2005; LaValle 2006). The control $u \in \mathcal{U}$ is generally selected uniformly at random to allow subsequent calls to extend the tree along new directions. Intermediate states along the trajectory are also added to the tree, as suggested in (Choset et al. 2005; LaValle 2006). The search terminates successfully when a vertex v_{new} is added to $\mathcal{T}$ such that $\text{map}_{\mathcal{S}\mapsto\mathcal{Q}}(v_{\text{new}}.s)$ satisfies ϕ_{goal}. The solution trajectory is then obtained by concatenating the collision-free and dynamically-feasible trajectories associated with the edges in $\mathcal{T}[E]$ connecting v_{init} to v_{new}.

Due to challenges posed by the high dimensionality of the continuous space $\mathcal{S}$, collision-avoidance requirements, differential constraints imposed by dynamics, and the complexity of the discrete specification, the success of the search depends on the ability of SamplSGD to effectively and selectively sample and explore $\mathcal{S}$. SamplSGD employs symbolic action planning to identify regions in $\mathcal{S}$ that sampling-based motion planning can then selectively sample and explore to significantly advance the search for a collision-free and dynamically-feasible trajectory that satisfies ϕ_{goal}.

In particular, SamplSGD groups the tree vertices according to the mapping function $\text{map}_{\mathcal{S}\mapsto\mathcal{Q}}$. Let Γ denote the list of all such groups, where $\Gamma_q \in \Gamma$ contains all the vertices $v \in \mathcal{T}$ such that $\text{map}_{\mathcal{S}\mapsto\mathcal{Q}}(v.s) = q$, i.e., each time a vertex v is added to $\mathcal{T}$, v is also added to $\Gamma_{\text{map}_{\mathcal{S}\mapsto\mathcal{Q}}(v.s)}$.

Consider one such group $\Gamma_q \in \Gamma$ and let a be an action whose precondition is satisfied by q. Let $a(q)$ denote the discrete state obtained as a result of a's effect on q. Sampling-based motion planning in SamplSGD, denoted by ExploreAction$(\mathcal{T}, \Gamma, \Gamma_q, a)$, can then advance the search by extending $\mathcal{T}$ from vertices associated with Γ_q (which satisfy a's precondition) toward the region of the continuous space $\mathcal{S}$ that satisfies a's postcondition, i.e., $\text{map}^{-1}_{\mathcal{S}\mapsto\mathcal{Q}}(a(q)) = \{s : s \in \mathcal{S} \wedge \text{map}_{\mathcal{S}\mapsto\mathcal{Q}}(s) = a(q)\}$. There is, however, an associated computational cost with each invocation of ExploreAction$(\mathcal{T}, \Gamma, \Gamma_q, a)$.

Algorithm 1 `SamplSGD`

Input: problem specification
$t_{\max}$: upper bound on computation time
Output: A collision-free and dynamically-feasible trajectory that satisfies the high-level goal specification or `null` if no solution is found

♢ *initialize data structures*
$\mathcal{T} \leftarrow \emptyset$; $\Gamma \leftarrow \emptyset$
$v_{\text{root}} \leftarrow$ new vertex; $v_{\text{root}}.s \leftarrow s_{\text{init}}$; v_{root}.parent $\leftarrow$ `null`
ADDVERTEX$(v_{\text{root}}, \mathcal{T}, \Gamma)$

♢ *core loop: interplay between symbolic action planning and sampling-based motion planning through action utilities*
while ELAPSEDTIME $< t_{\max} \wedge$ no solution path **do**
  $\Gamma_q \leftarrow$ SELECTGROUP(Γ)
  $a \leftarrow \Gamma_q$.curr_action
  status $\leftarrow$ EXPLOREACTION$(\mathcal{T}, \Gamma, \Gamma_q, a)$
  if status $=$ `solved` **then**
    $\zeta \leftarrow$ concatenate tree trajectories from root to last vertex
    return ζ
  UPDATEUTIL$(\mathcal{T}, \Gamma_q, a,$ status$)$
  Γ_q.curr_action $\leftarrow$ SELECTACTION(Γ_q)
  for each new $\Gamma_{q_{\text{new}}}$ added to Γ **do**
    $[a_i]_{i=1}^n \leftarrow$ SYMBOLICACTIONPLANNER$(\mathcal{A}, q_{\text{new}}, \phi_{\text{goal}})$
    $\Gamma_{q_{\text{new}}}$.actions $\leftarrow \{a_1\} \cup \Gamma_{q_{\text{new}}}$.actions
return `null`

This raises a central issue about which group $\Gamma_q \in \Gamma$ and which action a, among the many available options, should be selected for exploration at each invocation of EXPLOREACTION. To address this issue, `SamplSGD` maintains a running weight estimate

$$\text{UTIL}(\Gamma_q, a)$$

on the utility of having EXPLOREACTION$(\mathcal{T}, \Gamma_q, a)$ spend additional computational time attempting to extend $\mathcal{T}$ from vertices associated with Γ_q toward the region $\text{map}^{-1}_{\mathcal{S}\mapsto\mathcal{Q}}(a(q))$. `SamplSGD` updates UTIL$(\Gamma_q, a)$ based on new information gathered by sampling-based motion planning during each invocation of EXPLOREACTION. The objective of UTIL(Γ_q, a) is three-fold:

(i) give high utility to (Γ_q, a) when the action plan from the discrete state $a(q)$ to a discrete state that satisfies ϕ_{goal} is short. This is to bias the search in the continuous space $\mathcal{S}$ so that the sampling-based motion planner follows action plans that can quickly lead to a solution.

(ii) give high utility to (Γ_q, a) when it is under-explored, since additional exploration by the sampling-based motion planner could advance the search further.

(iii) give low utility to (Γ_q, a) when it is over-explored, since over-exploration does not bring much new information and wastes valuable computational time.

The core of `SamplSGD` interleaves symbolic action planning, sampling-based motion planning, and updates to utility estimates in order to effectively compute a collision-free and dynamically-feasible trajectory that satisfies ϕ_{goal}:

- Use symbolic action planning and the utility estimates to select a group $\Gamma_q \in \Gamma$ and an action a whose precondition is satisfied by q. Bias the selection process toward pairs (Γ_q, a) associated with high utilities.
- Use EXPLOREACTION$(\mathcal{T}, \Gamma_q, a)$ for a short period of time to extend $\mathcal{T}$ from vertices associated with Γ_q toward continuous states in $\text{map}^{-1}_{\mathcal{S}\mapsto\mathcal{Q}}(q)$, which satisfy a's post-condition.
- Update the utility estimates UTIL(Γ_q, a) based on new information gathered from EXPLOREACTION$(\mathcal{T}, \Gamma_q, a)$.

This interplay of symbolic action planning and sampling-based motion planning through the utility estimates is a crucial component of the computational efficiency of `SamplSGD`. In particular, it allows `SamplSGD` to make proper use of the computational time by selectively sampling and exploring those regions of the continuous space $\mathcal{S}$ that allow `SamplSGD` to significantly advance the search for a collision-free and dynamically-feasible trajectory that satisfies ϕ_{goal}. Pseudocode is given in Algo. 1. Details of the main components in `SamplSGD` follow.

Symbolic Action Planning

SYMBOLICACTIONPLANNER$(\mathcal{A}, q, \phi_{\text{goal}})$ computes an action plan $[a_i]_{i=1}^n$, which transforms the discrete state q to a discrete state that satisfies ϕ_{goal}. This is the only requirement imposed on SYMBOLICACTIONPLANNER, since `SamplSGD` uses it as a black-box. Therefore, `SamplSGD` can take advantage of research in AI and plug-in efficient symbolic actions planners (Blum and Furst 1997; Bonet and Geffner 2001; Keyder and Geffner 2008; Hoffmann and Nebel 2001; Hoffmann 2003; Ghallab, Nau, and Traverso 2004). These action planners, which reason about the discrete problem symbolically, are capable of effectively handling large discrete spaces and complex specifications.

Action Selection

Each $\Gamma_q \in \Gamma$ maintains a list of actions, written as Γ_q.actions, that are available for the selection process. This list is generally a small subset of all the actions whose preconditions are satisfied by q. For each $a \in \Gamma_q$.actions, Γ_q maintains a utility estimate, UTIL(Γ_q, a). An action $a \in \Gamma_q$.actions is then selected with probability proportional to its utility, i.e.,

$$\text{ProbSelect}_{\Gamma_q}(a) = \frac{\text{UTIL}(\Gamma_q a)}{\sum_{a' \in \Gamma_q.\text{actions}} \text{UTIL}(\Gamma_{q'}, a')}.$$

The selected action is kept as Γ_q.curr_action.

In this way, actions associated with high utilities are selected more often. This allows the sampling-based motion planner to quickly expand the search toward promising directions. At the same time, each available action has a non-zero probability of being selected, which is important to ensure that the search also expands along new directions.

When the group Γ_q is first created, the function SYMBOLICACTIONPLANNER$(\mathcal{A}, q, \phi_{\text{goal}})$ is invoked to compute an action plan $[a_i]_{i=1}^n$, which transforms the discrete state q to a discrete state that satisfies ϕ_{goal}. Since the overall objective is to satisfy ϕ_{goal}, if there are no action plans from q that satisfy ϕ_{goal}, then Γ_q is deleted from

Γ. Otherwise, the first action of the plan, a_1, is added to $\Gamma_q.\text{actions}$. Thus, initially, $\Gamma_q.\text{actions}$ contains only one action. Note that this provides an opportunity to use symbolic action planners that can efficiently compute the first action of an action plan.

As the search progresses and new information is gathered by the sampling-based motion planner, the utilities of actions in $\Gamma_q.\text{actions}$ are updated to take into account this new information. When the utilities of all the actions in $\Gamma_q.\text{actions}$ fall below a certain threshold, SYMBOLICACTIONPLANNER$(\mathcal{A}, q, \phi_{\text{goal}})$ is invoked again to compute a new action plan $[a_i]_{i=1}^n$, where $a_1 \notin \Gamma_q.\text{actions}$. If it succeeds, as during initialization, the first action of the plan is added to $\Gamma_q.\text{actions}$ and is also made available for selection.

In this way, the search is made broader when it becomes difficult to make significant progress using the current actions to guide the sampling-based motion planner. This allows the sampling-based motion planner to explore new directions, which could lead to further progress in the search for a collision-free a dynamically-feasible trajectory that satisfies the goal specification.

Group Selection

The utility of a group $\Gamma_q \in \Gamma$ is defined as the utility of the action currently selected in Γ_q, i.e.,

$$\text{UTIL}(\Gamma_q) = \text{UTIL}(\Gamma_q, \Gamma_q.\text{curr\_action}).$$

Then, as during action selection, a group Γ_q is selected from Γ with probability proportional to its utility, i.e.,

$$\text{ProbSelect}_\Gamma(\Gamma_q) = \frac{\text{UTIL}(\Gamma_q)}{\sum_{\Gamma_{q'} \in \Gamma} \text{UTIL}(\Gamma_{q'})},$$

which aims to strike a balance between being greedy and being methodical by giving preference to groups associated with high utilities, without ignoring other groups in Γ.

Action Utility

Drawing from earlier work in sampling-based motion planning (Burns and Brock 2007; Sánchez and Latombe 2002; Ladd and Kavraki 2005; Plaku, Kavraki, and Vardi 2007; 2008b), the utility estimates proposed in this paper are designed to be computationally efficient and to work well in practice. Further improving the proposed utility estimates remains an important direction for future research.

$\text{UTIL}(\Gamma_q, a)$ is initialized based on the length of the action plan $[a_i]_{i=1}^n$ (with $a = a_1$) computed by SYMBOLICACTIONPLANNER$(\mathcal{A}, q, \phi_{\text{goal}})$ when Γ_q is first created. More specifically,

$$\text{UTIL}(\Gamma_q, a) \leftarrow \frac{1}{|\text{plan}(a, q)|^2}.$$

Since the overall objective is to compute a solution as quickly as possible, this initialization assigns high utility to (Γ_q, a_1) when the action plan is short.

`SamplSGD` uses ideas from reinforcement learning to update $\text{UTIL}(\Gamma_q, a)$ based on information gathered by the sampling-based motion planner. After each invocation of EXPLOREACTION$(\mathcal{T}, \Gamma, \Gamma_q, a)$, the utility estimate $\text{UTIL}(\Gamma_q, a)$ is updated a as follows:

$$\text{UTIL}(\Gamma_q, a) \leftarrow \text{UTIL}(\Gamma_q, a)(1 - \alpha(\Gamma_q, a)) +$$
$$\alpha(\Gamma_q, a)(r(\Gamma_q, a) + \gamma \max_{a' \in \Gamma_{a(q)}.\text{actions}} \text{UTIL}(\Gamma_{a(q)}, a')).$$

The learning rate, which determines the extent at which the new information will affect the utility value, is set to $\alpha(\Gamma_q, a) = \frac{1}{1+\sqrt{n_{\text{sel}}(\Gamma_q, a)}}$, where $n_{\text{sel}}(\Gamma_q, a)$ denotes the number of times $a \in \Gamma_q.\text{actions}$ has been selected.

The discount factor, which determines the importance of future rewards, is set to a constant value $\gamma = 1/16$. This allows the utility function to be greedy toward current rewards, while also striving for long-term high rewards.

The reward $r(\Gamma_q, a)$ is based on the success the sampling-based motion planner has had in extending $\mathcal{T}$ to $\text{map}_{\mathcal{S} \mapsto \mathcal{Q}}^{-1}(a(q))$. In particular, $r(\Gamma_q, a)$ depends on the number of tree vertices in $\text{map}_{\mathcal{S} \mapsto \mathcal{Q}}^{-1}(a(q))$. More specifically, it is computed as

$$r(\Gamma_q, a) = \begin{cases} 0, & \text{if } |\Gamma_{a(q)}| = 0 \\ 1/|\Gamma_{a(q)}|, & \text{otherwise.} \end{cases}$$

In this way, the reward is zero when the sampling-based motion planner has not yet extended $\mathcal{T}$ to the region $\text{map}_{\mathcal{S} \mapsto \mathcal{Q}}^{-1}(a(q))$. When $\mathcal{T}$ reaches $\text{map}_{\mathcal{S} \mapsto \mathcal{Q}}^{-1}(a(q))$, the reward becomes high, since now the search has advanced further. As more and more tree vertices are added to $\Gamma_{a(q)}$, the region starts becoming over-explored, and, consequently, the reward is reduced.

Sampling-based Motion Planning

The objective of EXPLOREACTION$(\mathcal{T}, \Gamma, \Gamma_q, a)$ is to extend $\mathcal{T}$ toward the region $\text{map}_{\mathcal{S} \mapsto \mathcal{Q}}^{-1}(a(q))$, so that $\mathcal{T}$ can follow in the continuous space $\mathcal{S}$ the discrete action a. EXPLOREACTION proceeds in an iterative fashion, as illustrated in Algo. 2. At each iteration, EXPLOREACTION first selects a vertex v from which to extend $\mathcal{T}$ (Algo. 2:2). At a second step, EXPLOREACTION generates a dynamically-feasible trajectory $\zeta : [0, T] \rightarrow \mathcal{S}$ that starts at $v.s$, i.e., $\zeta(0) = v.s$ (Algo. 2:3). ζ is generated by sampling a control $u \in \mathcal{U}$ uniformly at random and simulating the dynamics forward in time starting from $v.s$ until a collision occurs, a state-constraint is violated, or a maximum number of steps is exceeded (Choset et al. 2005; LaValle 2006) (Algo. 2:4–11). Intermediate collision-free states along ζ are added as new vertices to $\mathcal{T}$ (Algo. 2:9). If a new vertex, v_{new}, satisfies the goal specification (i.e., $\text{map}_{\mathcal{S} \mapsto \mathcal{Q}}(v_{\text{new}})$ satisfies ϕ_{goal}), then a solution trajectory is obtained by concatenating the tree trajectories from the root of $\mathcal{T}$ to v_{new} (Algo. 2:10–11). Otherwise, the above steps are repeated several times. The remainder of the section describes in more detail the vertex-selection and trajectory-generation strategies.

Vertex Selection: Over the years, numerous vertex-selection strategies have been proposed in motion-planning literature that rely on distance metrics, nearest neighbors,

Algorithm 2 EXPLOREACTION$(\mathcal{T}, \Gamma, \Gamma_q, a)$

Input: $\mathcal{P}$: problem specification
$\mathcal{T}$: search tree
Γ: mapping of search tree over the discrete space
Γ_q: selected group
a: selected action

Output: A collision-free and dynamically-feasible trajectory that satisfies the high-level goal specification or `null` if no solution is found

for several times **do**
  $v \leftarrow$ SELECTVERTEX$(\mathcal{T}, \Gamma_q)$
  $\zeta \leftarrow$ GENERATETRAJ(v, a)
  parent $\leftarrow v$
  status $\leftarrow$ `ok`
  for $t = \delta; t \leq |\zeta| \wedge$ status $\neq$ `rejected`; $t \leftarrow t + \delta$ **do**
    $s_t \leftarrow \zeta(t)$
    if s_t is valid, e.g., collision-free **then**
      $v_{\text{new}}.[s, \text{parent}] \leftarrow$ new vertex $[s_t, \text{parent}]$
      parent $\leftarrow v_{\text{new}}$
      ADDVERTEX$(\mathcal{T}, \Gamma, v_{\text{new}}, a)$
      if $\text{map}_{\mathcal{S} \mapsto \mathcal{Q}}(v_{\text{new}}.s)$ satisfies $\mathcal{P}.\phi_{\text{goal}}$ **then**
        return TRAJ$(\mathcal{T}, v_{\text{new}})$
    else
      status $\leftarrow$ `rejected`
return `null`

probability distributions, and many others, as surveyed in (Choset et al. 2005; LaValle 2006). Drawing from this body of research and earlier work (Plaku, Kavraki, and Vardi 2007; 2008b), the vertex-selection strategy in this paper combines the advantages of several successful techniques.

S1: To bias the growth of $\mathcal{T}$ toward $\text{map}^{-1}_{\mathcal{S} \mapsto \mathcal{Q}}(a(q))$, one approach, proposed in (LaValle and Kuffner 2001), that has been shown to work well in practice is to first sample a continuous state, s, such that $\text{map}_{\mathcal{S} \mapsto \mathcal{Q}}(s) = a(q)$. The vertex from which to extend $\mathcal{T}$ is then selected from the vertices in $\Gamma_q \cup \Gamma_{a(q)}$ as the vertex whose associated continuous state, $v.s$, is the closest to s according to a distance metric. The effect of this strategy is to pull $\mathcal{T}$ toward $\text{map}^{-1}_{\mathcal{S} \mapsto \mathcal{Q}}(a(q))$. To improve the computational efficiency, the distance metric in this paper is defined over a low-dimensional projection (e.g., by considering only the position component). Moreover, approximate nearest neighbors (Plaku and Kavraki 2006; 2007) are used to speed up computation without any significant loss in accuracy.

S2: Another objective of EXPLOREACTION is to grow $\mathcal{T}$ toward unexplored or sparsely explored areas. This avoids over-exploration and leads the search toward new directions. As proposed in (Sánchez and Latombe 2002; Ladd and Kavraki 2005; Plaku, Kavraki, and Vardi 2007; 2008b), an effective strategy for these purposes is to further group the vertices in $\Gamma_q \cup \Gamma_{a(q)}$ into cells based on a low-dimensional projection of the continuous space $\mathcal{S}$. A vertex from which to extend $\mathcal{T}$ is then obtained by first selecting a cell c and then selecting a vertex from c uniformly at random. The cell c could be selected uniformly at random (Sánchez and Latombe 2002) or based on coverage estimates (Ladd and Kavraki 2005; Plaku, Kavraki, and Vardi 2007; 2008b). The effect of this strategy is to push $\mathcal{T}$ toward unexplored or sparsely explored areas.

S3: A third objective of EXPLOREACTION is to further extend the search forward by increasing the depth of $\mathcal{T}$, similar to depth-first search in a discrete setting. The vertex from which to extend $\mathcal{T}$ is then selected uniformly at random from the last vertices (around 20 is shown to work well in practice) added to $\mathcal{T}$ as a result of previous invocations of EXPLOREACTION$(\mathcal{T}, \Gamma, \Gamma_q, a)$.

The overall vertex-selection strategy then simply selects at each iteration one of the above strategies uniformly at random. The effect is a vertex-selection strategy that pulls $\mathcal{T}$ toward $\text{map}^{-1}_{\mathcal{S} \mapsto \mathcal{Q}}(a(q))$, while avoiding over-exploration, finding new directions, increasing sampling in sparsely explored areas, and expanding the search depth.

Trajectory Generation: The trajectory-generation strategy discussed in this section is designed to work well in practice for a wide class of systems and actions. It is possible, however, to further improve these strategies by taking advantage of the problem specification. In particular, one can design specific trajectory-generation strategies for each action a. As an example, if the action is "GraspObject," then the trajectory-generation strategy can be designed to produce open-and-closing motions of the end-effector tool. In this way, action-specific strategies can be used to further improve the overall effectiveness of EXPLOREACTION$(\mathcal{T}, \Gamma, \Gamma_q, a)$.

4. Experiments and Results

The proposed method is tested on a object-manipulation task. As illustrated in Fig. 1, the high-level specification requires the car to pickup the objects and transfer each object to its corresponding area as indicated by the object label, i.e., transfer object i to area i ($i = 1, 2, 3, 4$). The car picks up an object by touching it, but it cannot pick up more than one object at a time. The car can temporarily release and pickup up objects at any location. The car should avoid collisions with the objects at all times (except the object that it picks up, which is allowed to be in contact with the car). No collisions should occur between two objects.

The car is modeled as a second-order dynamical systems. Details can be found in (LaValle 2006, pp. 744). The state $s = (x, y, \theta, v, \psi)$ consists of the position $(x, y) \in \mathbb{R}^2$ ($|x|, |y| \leq 3.75m$), orientation $\theta \in [-\pi, \pi)$, velocity v ($|v| \leq 3m/s$), and steering-wheel angle ψ ($|\psi| \leq 50°$). The car is controlled by setting the acceleration u_0 ($|u_0| \leq 1m/s^2$) and the rotational velocity of the steering-wheel angle u_1 ($|u_1| \leq 100°/s$). The equations of motions are $\dot{x} = v\cos(\theta)$; $\dot{y} = v\sin(\theta)$; $\dot{\theta} = v\tan(\psi)/L$; $\dot{v} = u_0$; $\dot{\psi} = u_1$, where $L = 0.5m$ is the distance between the front and rear axles. The body length and width are set to L and $0.5L$, respectively. The scaling factor is $1m = 0.14$ workspace units.

This object-manipulation task provides several challenges. The discrete space is complex. There is a large number of action plans, which provide solutions in the discrete setting. Due to the arrangement of objects and obstacles, the objects cannot be transferred just in any order. As an example, before transferring object 1 to area 1, the car must

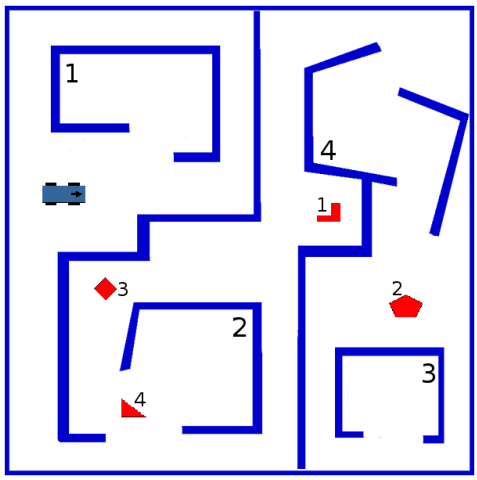

Figure 1: Object-manipulation task used in the experiments. The car (with second-order dynamics) needs to pick up the objects (shown in red) one at a time and place them in the areas corresponding to their labels, i.e., object i should be placed in area i. Obstacles are shown in blue.

transfer objects 3 and 4 to some other areas to make room for object 1. The small openings and the second-order dynamics make it difficult for the car to wiggle its way through as it transfers objects from one location to another.

The experiments provide a promising initial validation of the proposed method, SamplSGD. These experiments highlight the importance of the interplay between symbolic action planning and sampling-based motion planning. In fact, SamplSGD is significantly more efficient than state-of-the-art sampling-based motion planners, such as RRT(LaValle and Kuffner 2001), ADDRRT(Jaillet et al. 2005), EST(Hsu et al. 2002). Without symbolic action planning, it is nearly impossible for sampling-based motion planning just by itself to find a solution, since the solution trajectory must satisfy complex high-level specifications. We conducted numerous trials (20 per planner). In each case these sampling-based path planners failed to find a solution, even though the time limit was set to 2 hours per trial. On the other hand, SamplSGD was able to efficiently find a solution in a matter of a few minutes (average 6 mins).

5. Discussion

This paper proposed a multi-layered approach, SamplSGD, which incorporates symbolic, geometric, and differential constraints into sampling-based motion planning. Given a high-level goal specification in a planning-domain definition language, such as STRIPS, SamplSGD computes a collision-free and dynamically-feasible trajectory that satisfies the goal specification. The crucial component of SamplSGD is the interplay between symbolic action planning and sampling-based motion planning. SamplSGD leverages from state-of-the-art sampling-based motion planning the underlying idea of searching for a solution by selectively sampling and exploring the continuous space. To effectively incorporate collision-avoidance and differential constraints imposed by underlying dynamics, SamplSGD conducts a tree-based exploration of the continuous state space. Drawing from research in AI, SamplSGD uses symbolic action planning in novel ways to identify discrete actions and regions of the continuous space that sampling-based motion planning can selectively sample and explore to significantly advance the search. The planning layers interact with each-other through estimates on the utility of each action, which are computed by reinforcement-learning techniques based on information gathered by the sampling-based motion planner. Simulation experiments on an object-manipulation task with a second-order dynamical model provide promising initial validation.

An objective for future work is to further improve the interplay between the different components in SamplSGD. We are also working on adapting the proposed approach to real robotic platforms. This will allow us to tackle increasingly complex problems arising in robot manipulation and automation.

References

Alami, R.; Laumond, J.; and Siméon, T. 1995. Two manipulation planning algorithms. In *Int. Work. Algo. Found. Robot.* 109–125.

Amato, N. M.; Bayazit, B.; Dale, L.; Jones, C.; and Vallejo, D. 1998. OBPRM: An obstacle-based PRM for 3d workspaces. In *Int. Work. Algo. Found. Robot.* 156–168.

Arkin, R. C. 1990. Integrating behavioral, perceptual and world knowledge in reactive navigation. *Robotics and Autonomous Systems* 6:105–122.

Baier, C., and Katoen, J. P. 2008. Principles of model checking.

Belta, C.; Bicchi, A.; Egerstedt, M.; Frazzoli, E.; Klavins, E.; and Pappas, G. J. 2007. Symbolic planning and control of robot motion. *IEEE Robotics and Automation Magazine* 14(1):61–71.

Blum, A. L., and Furst, M. L. 1997. Fast planning through planning graph analysis. *Artificial Intelligence* 90(1-2):281–300.

Bonet, B., and Geffner, H. 2001. Planning as heuristic search. *Artificial Intelligence* 129(1-2):5–33.

Burns, B., and Brock, O. 2007. Single-query motion planning with utility-guided random trees. In *IEEE Int. Conf. Robot. Autom.*, 3307–3312.

Cambon, S.; Alami, R.; and Gravot, F. 2009. A hybrid approach to intricate motion, manipulation and task planning. *Int. J. Robot. Res.* (1):104–126.

Cambon, S.; Gravot, F.; and Alami, R. 2004. A robot task planner that merges symbolic and geometric reasoning. In *European Conf. on Artificial Intelligence*, volume 16, 895–899.

Choset, H.; Lynch, K. M.; Hutchinson, S.; Kantor, G.; Burgard, W.; Kavraki, L. E.; and Thrun, S. 2005. *Principles of Robot Motion: Theory, Algorithms, and Implementations*. MIT Press.

Fainekos, G. E.; Girard, A.; Kress-Gazit, H.; and Pappas, G. J. 2007. Temporal logic motion planning for dynamic mobile robots. *Automatica* 45(2):343–352.

Fikes, R., and Nilsson, N. 1971. STRIPS: a new approach to the application of theorem proving to problem solving. *Artificial Intelligence* 2(3-4):189–208.

Ghallab, M.; Howe, A.; Knoblock, C.; McDermott, C.; Ram, A.; Veloso, M. Weld, D.; and Wilkins, D. 1998. PDDL–the planning domain definition language.

Ghallab, M.; Nau, D. S.; and Traverso, P. 2004. *Automated Planning: theory and practice*. Morgan Kaufmann.

Gravot, F.; Cambon, S.; and Alami, R. 2003. aSyMov: a planner that deals with intricate symbolic and geometric problems. In *Int. Symp. Robotics Research*, volume 15 of *Springer Tracts in Advanced Robotics*. 100–110.

Hoffmann, J., and Nebel, B. 2001. The FF planning system: Fast plan generation through heuristic search. *J. of Artificial Intelligence Research* 14:253–302.

Hoffmann, J. 2003. The Metric-FF planning system: Translating "ignoring delete lists" to numeric state variables. *J. of Artificial Intelligence Research* 20(20):291–341.

Hsu, D.; Kindel, R.; Latombe, J.-C.; and Rock, S. 2002. Randomized kinodynamic motion planning with moving obstacles. *Int. J. Robot. Res.* 21(3):233–255.

Jaillet, L.; Yershova, A.; LaValle, S. M.; and Simeon, T. 2005. Adaptive tuning of the sampling domain for dynamic-domain RRTs. In *IEEE/RSJ Int. Conf. Intell. Robot. Syst.*, 4086–4091.

Kavraki, L. E.; Švestka, P.; Latombe, J.-C.; and Overmars, M. H. 1996. Probabilistic roadmaps for path planning in high-dimensional configuration spaces. *IEEE Trans. Robot. Autom.* 12(4):566–580.

Keller, H. 1992. *Numerical Methods for Two-Point Boundary-Value Problems*. New York, NY: Dover.

Keyder, E., and Geffner, H. 2008. Heuristics for planning with action costs revisited. In *European Conf. on Artificial Intelligence*, 588–592.

Ladd, A. M., and Kavraki, L. E. 2005. Motion planning in the presence of drift, underactuation and discrete system changes. In *Robotics: Sci. and Systems*, 233–241.

LaValle, S. M., and Kuffner, J. J. 2001. Randomized kinodynamic planning. *Int. J. Robot. Res.* 20(5):378–400.

LaValle, S. M. 2006. *Planning Algorithms*. Cambridge, MA: Cambridge University Press.

Marthi, B.; Russell, S.; and Wolfe, J. 2007. Angelic semantics for high-level actions. In *Intl Conf on Automated Planning and Scheduling*.

Nielsen, C., and Kavraki, L. E. 2000. A two level fuzzy PRM for manipulation planning. In *IEEE/RSJ Int. Conf. Intell. Robot. Syst.*, 1716–1722.

Payton, D. W.; Rosenblatt, J. K.; and Keirsey, D. M. 1990. Plan guided reaction. *IEEE Trans. on Systems, Man, and Cybernetics* 20(6):1370–1372.

Pednault, E. P. D. 1994. ADL and the state-transition model of action. *J Logic Computation* 4(5):467–512.

Plaku, E., and Kavraki, L. E. 2006. Quantitative analysis of nearest-neighbors search in high-dimensional sampling-based motion planning. In *Int. Work. Algo. Found. Robot.*, Springer Tracts in Advanced Robotics.

Plaku, E., and Kavraki, L. E. 2007. Nonlinear dimensionality reduction using approximate nearest neighbors. In *SIAM Int.Conf. on Data Mining*, 3711–3716.

Plaku, E.; Kavraki, L. E.; and Vardi, M. Y. 2007. Discrete search leading continuous exploration for kinodynamic motion planning. In *Robotics: Sci. and Systems*, 326–333.

Plaku, E.; Kavraki, L. E.; and Vardi, M. Y. 2008a. Hybrid systems: From verification to falsification by combining motion planning and discrete search. *Formal Methods in System Design*.

Plaku, E.; Kavraki, L. E.; and Vardi, M. Y. 2008b. Impact of workspace decompositions on discrete search leading continuous exploration (DSLX) motion planning. In *IEEE Int. Conf. Robot. Autom.*, 3751–3756.

Plaku, E.; Kavraki, L. E.; and Vardi, M. Y. 2009. Falsification of LTL safety properties in hybrid systems. In *Intl Conf on Tools and Algorithms for the Construction and Analysis of Systems*.

Russell, S., and Norvig, P. 2002. *Artifical Intelligence: A Modern Approach*. Prentice Hall, 2 edition.

Saffiotti, A.; Konolige, K.; and Ruspini, E. H. 1995. A multivalued logic approach to integrating planning and control. *Artificial Intelligence* 76(1-2):481–526.

Sánchez, G., and Latombe, J.-C. 2002. On delaying collision checking in PRM planning: Application to multi-robot coordination. *Int. J. Robot. Res.* 21(1):5–26.