Task-Motion Planning for Safe and Efficient Urban Driving

Yan Ding, Xiaohan Zhang, Xingyue Zhan, Shiqi Zhang

Abstract—Autonomous vehicles need to plan at the task level to compute a sequence of symbolic actions, such as merging left and turning right, to fulfill people’s service requests, where efficiency is the main concern. At the same time, the vehicles must compute continuous trajectories to perform actions at the motion level, where safety is the most important. Task-motion planning in autonomous driving faces the problem of maximizing task-level efficiency while ensuring motion-level safety. To this end, we develop algorithm Task-Motion Planning for Urban Driving (TMPUD) that, for the first time, enables the task and motion planners to communicate about the safety level of driving behaviors. TMPUD has been evaluated using a realistic urban driving simulation platform. Results suggest that TMPUD performs significantly better than competitive baselines from the literature in efficiency, while ensuring the safety of driving behaviors.

I. INTRODUCTION

Autonomous driving technologies have the great potential of reshaping urban mobility in people’s daily life [1], [2], [3]. To be deemed useful, autonomous vehicles must be time-efficient in accomplishing service tasks, which frequently requires symbolic actions such as “Merge left, go straight, turn left, and park right”, while at the same time ensuring safety in executing such actions on the road [4], [5], [6].

Generally, autonomous vehicles need to plan at the task level to compute a sequence of symbolic actions toward fulfilling service requests from people. In this process, how the actions are implemented in the real world is out of consideration at the task level. At the same time, vehicles must plan at the motion level to compute continuous trajectories, and desired control signals (e.g., for steering, accelerating, and braking) to implement the symbolic actions. While the task planner hopes that all the symbolic actions can be implemented by the vehicles, there is the safety concern that must be considered at the motion level. For instance, lane-changing behaviors can be dangerous in heavy traffic. Fig. 1 shows two situations that are dangerous (Left) and safe (Right) to a vehicle, respectively.

Although task planning (frequently referred to as behavior planning in autonomous driving [7]) and motion planning have been individually conducted in autonomous driving, there is little research from the literature focusing on the interaction between task and motion levels. There is the critical need of developing algorithms to bridge the gap between task planning and motion planning to help vehicles improve the task-completion efficiency while ensuring the safety of driving behaviors.

The robotics community has studied the integration of task and motion planning, mostly in manipulation domains [8], [9], [10], [11]. In comparison to those domains, autonomous driving algorithms must consider the uncertainty from the ego vehicle, and the surrounding objects (including other vehicles) on the road. The uncertainty must be quantitatively evaluated at the motion level, and taken into consideration for planning at the task level. For instance, when the left lane is busy and missing the next crossing does not introduce much extra distance, the task planner should avoid forcing the vehicle to merge left. Such behaviors are possible, only if the interactions between task and motion levels are enabled.

In this paper, we develop Task-Motion Planning for Urban Driving (TMPUD) for efficient and safe autonomous urban driving. TMPUD, for the first time, enables the interaction between task and motion planners through enabling the motion-level safety estimation and task-level replanning capabilities. The contribution of this research is twofold, including the new safety estimator, and the TMPUD algorithm. We have implemented and evaluated TMPUD using CARLA, an autonomous driving platform for simulating urban driving scenarios [12]. Results suggest TMPUD improves both safety and efficiency, in comparison to two baseline methods from the literature [13], [8].

II. RELATED WORK

We summarize three research areas that are the most relevant to this research, namely motion planning in autonomous driving, task planning in autonomous driving, and the integration of task and motion planning.

Motion-Level Planning for Autonomous Driving: Safety is the most importance at the motion level, and highly relies on the motion-level controllers. Early research in robotics
Task Planning for Autonomous Driving: Task planning has been applied to autonomous driving. For instance, one of the earliest works on this topic demonstrated that task planning techniques enable vehicles to complete complex tasks, such as to avoid temporary roadblocks [13] (we use this approach as a baseline in experiments). However, their work did not consider costs of driving behaviors, and hence performed poorly in task-completion efficiency. Similarly, task planners in [22], [23], [24] have no interaction with the motion planner, and safety was not modeled in generating the driving behaviors.

Moving forward, more recent research has enabled vehicles to periodically verify the task sequences and motion trajectories against the actual traffic situation [25]. In case of possible dangers detected at the motion level, re-planning is triggered at the task level. The main limitation of their work is that the triggering is deterministic, and highly depends on a safety threshold. The threshold must be set beforehand to ensure safety, which frequently produces over-conservative behaviors, and significantly reduces the task-level efficiency.

Task and Motion Planning: Researchers have integrated task and motion planning in robotics, where the primary domain is robot manipulation [26], [27], [28], [29], [8]. Research on manipulation is mostly concerned with the motion-level feasibility, e.g., in grasping and ungrasping behaviors, and accomplishing high-level tasks, such as stacking objects. Those methods did not consider the uncertainty from other agents (e.g., vehicles on the road). As a result, their systems produce over-optimistic (and hence risky) behaviors, assuming no other agents making changes in the world, and are not applicable to autonomous driving domains.

A journal paper has surveyed frameworks for autonomous driving [7], including works that plan at both task and motion levels. However, their motion planners do not provide any feedback to the task level, except for infeasible actions. In comparison, our TMPUD algorithm supports motion-level safety evaluation, and enables the task planner to dynamically adjust high-level plans to account for current road conditions toward accomplishing long-term driving tasks.

III. BACKGROUND

We very briefly summarize task planning and motion planning, the two building blocks of this research.

Task Planning: A task planning domain is specified by $D_t$, including a set of states, $S$, and a set of actions, $A$. We assume a factored state space such that each state $s \in S$ is defined by the values of a fixed set of variables; each action $a \in A$ is defined by its preconditions and effects. A utility function maps the state transition to a real number, which takes both cost function $Cost((s, a, s'))$ and safety function $Safe((s, a, s'))$ into account. Specifically, the cost and safety functions respectively reflect the cost and safety of conducting action $a$ in state $s$.

Given domain $D_t$ and a task planning problem, we want to compute a plan $p \in P$, starting from an initial state $s^{init} \in S$ and finishing in a goal state $s^g \in S$. A plan $p$ consists of a sequence of transitions that can be represented as: $p =$
\( \langle s_0, a_0, \ldots, s_{N-1}, a_{N-1}, s_N \rangle \), where \( s_0 = s^{\text{init}} \), \( s_N = s^g \) and \( P \) denotes a set of satisfactory plans. Task planner \( P^t \) can produce an optimal plan \( p^* \) among all satisfactory plans, where \( \gamma \) is a constant coefficient and \( \gamma > 0 \).

\[ p^* = \arg \min_{p \in P} \sum_{\langle s, a, s' \rangle \in p} \left[ \text{Cost}(\langle s, a, s' \rangle) + \frac{\gamma}{1 + e^{\text{Safe}(\langle s, a, s' \rangle) - 1}} \right] \]

**Motion Planning:** A motion planning domain is specified by \( D^m \), where we directly search in 2D space constrained by the urban road network. Some parts of the space are designated as free space, and the rest are designated as obstacles. The 2D space is represented as a region in Cartesian space such that the position and orientation of the vehicle can be uniquely represented as a pose, denoted by \( x \).

Given domain \( D^m \), a motion planning problem can be solved by an initial pose \( x^i \) and a goal pose \( x^g \). The motion planning problem is solved by a motion planner \( P^m \) consisting of path planner and tracking planner into two phases. In the first one, a path planner computes a collision-free trajectory \( \xi \) connecting pose \( x^i \) and pose \( x^g \) taking into account any motion constraints on the part of the vehicle with minimal trajectory length. In the second one, a tracking controller computes desired control signals to drive the vehicle to follow the computed trajectory. Due to the fundamental difference between representations at task and motion levels, in line with past research [26], [29], [8], [11], we use a state mapping function, \( f : X = f(s) \), to map the symbolic state \( s \) into a set of feasible poses \( X \) in continuous space, for motion planner to sample from. We assume the availability of at least one pose \( x \in X \) in each state \( s \), such that the vehicle is in the free space of \( D^m \). If it is not the case, the state \( s \) is declared infeasible.

**Algorithm 1 Safety Estimation**

**Input:** Symbolic action \( \langle s, a, s' \rangle \), state mapping function \( f \), motion planner \( P^m \), control operation sets \( \Delta \) and \( \Theta \)

1: Sample initial and goal poses, \( x \leftarrow f(s) \) and \( x' \leftarrow f(s') \), given action \( \langle s, a, s' \rangle \), and \( f \).
2: Compute a collision-free trajectory, \( \xi \), using \( P^m \), where \( \xi(t_1) = x, \xi(t_2) = x' \), and \([t_1, t_2]\) is the horizon
3: Predict trajectory \( \xi(t) \) for the \( i \)th surrounding vehicle, where \( i \in [1, \ldots, N] \), and \([t_1, t_2]\) is the horizon
4: while each vehicle \( V_i \) do
5: Compute safe control set \( U_S(t) \) between the ego vehicle and vehicle \( V_i \) at time \( t \in [t_1, t_2] \), where \( U_S(t) \subset \Delta \times \Theta \) and \( t = t_1 + \omega \times i, i \leq \lceil \frac{t_2 - t_1}{\omega} \rceil \)
6: Sample \( M \) elements \( \langle \delta, \theta \rangle \) randomly from set \( \Delta \times \Theta \) and compute the probability \( o(t) \) of the elements falling in set \( U_S(t) \)
7: Convert a list of estimated safety values, \( \{o_i(t)\} \), into a scalar \( o_i^* \) using Eqn. 1
8: end while
9: return \( \min\{o_i^*, i = 1, \ldots, N\} \)

**Safety Estimation Algorithm:** Algorithm 1 summarizes the procedure of our safety estimation algorithm. The input includes symbolic action \( \langle s, a, s' \rangle \), state mapping function \( f \), motion planner \( P^m \) consisting of path planner and tracking controller, and the controller’s operation specification sets \( \Delta \) and \( \Theta \). The output is the estimated safety value \( \text{Safe}(\langle s, a, s' \rangle) \in [0.0, 1.0] \).

Lines 1-3 aim to obtain the short-period trajectories of the ego and surrounding vehicles, where \( V_i, i \in [1, \ldots, N] \), is the \( i \)th vehicle within the ego vehicle’s sensing range. More specifically, we first sample a pair of feasible initial and goal poses for the symbolic actions using the state mapping function (Line 1). Taking these two poses as input, the motion planner then computes a continuous trajectory for our ego vehicle for a short period of time \([t_1, t_2]\) (Line 2), where \( t_1 \) is the current time, and \( t_2 = t_1 + T \) indicates the time horizon of the ego vehicle. We predicate surrounding vehicles’ trajectories, assuming their linear and angular speeds being stationary (Line 3), though there are more advanced methods [30], [31], which is beyond the scope of this research.

Lines 4-8 present a control loop that computes the safety estimation between the ego vehicle and the surrounding vehicles \( V_i \), where \( i \in [1, \ldots, N] \), given that the ego vehicle is performing action \( \langle s, a, s' \rangle \) at the motion level. We compute a safe control set \( U_S(t) \), similar to [19], that includes all safe control signals with regard to vehicle \( V_i \) at time \( t \) (Line 5). Parameter \( \omega \) controls the sampling interval.
In Line 6, we randomly sample $M$ elements from the set $\Delta \times \Theta$, and compute probability $o_i(t)$ of the sampled elements falling in set $U^i(t)$. We convert a list of values of safety estimation $\{o_i(t)\}$ into a single value $o^*_i$ using eqn 1, where $\max$ and $\text{mean}$ are two functions to calculate the maximum and mean value of a list, respectively (Line 7).

Although all surrounding vehicles can potentially introduce risks to the ego vehicle, we assume the ego vehicle only considers the most dangerous vehicle. Accordingly, Line 9 is used for selecting the minimum value, $o^*_i$, $i \in [1, \cdots, N]$, as the overall safety value:

$$o^*_i = \frac{\max_{t \in T} \{o_i(t)\} + \text{mean}_{t \in T} \{o_i(t)\}}{2} \quad (1)$$

where $T = t_1 + \omega \times i, \ 0 \leq i \leq \frac{(t_2-t_1)}{\omega}$

B. TMPUD

Our motion planner $P^m$ computes both costs (trajectory lengths) and safety values of the ego vehicle’s navigation actions, which have been discussed in Section IV-A. Here, we focus on the main contribution of this work on enabling interactive task-motion planning for urban driving.

Terminology: We use $s^{\text{init}}$ to represent the initial state of the ego vehicle, and the goal (service request from people) is specified using $s^g$. Our task planner $P^t$ computes a sequence of symbolic actions, and it requires two functions that are initialized and updated within the algorithm, including cost function $Cost$, and safety estimation function $Safe$. Motion planner $P^m$ is used for computing motion trajectories, and generating control signals to move the ego vehicle. The state mapping function $f$ is used for mapping symbolic states to 2D coordinates in continuous spaces.

The TMPUD Algorithm: Algorithm 2 summarizes the procedure of TMPUD. It starts by initializing the cost and safety estimation functions (Lines 1 and 2). Cost function $Cost$ is initialized using A star algorithm provided by CARLA, as shown in Line 1. In Line 2, TMPUD optimistically initializes the safety estimation function by setting 1.0 to all actions, indicating all task-level actions are completely safe. After that, an optimal task plan, $p^* = (s^{\text{init}}, a_0, s_1, \cdots, s^g)$, is computed in Line 3. The head and tail elements of the plan, $s^{\text{init}}$ and $s^g$, correspond to the initial and goal poses respectively.

Lines 4-19 form TMPUD’s main control loop that enables the interaction between task and motion planners. The loop’s termination condition is the task-level plan being empty, i.e., the goal has been achieved (Line 4). Specifically, TMPUD estimates the safety level, $\mu$, of action $\langle s, a, s' \rangle$ (Line 5). Functions $Safe$ and $Cost$ are updated using $\mu$ and A* search in Line 6. Then a new optimal plan $p'$ is computed in Line 7. Lines 8-18 is for plan monitoring and action execution. If the task planner suggests the same plan (Line 8), the vehicle will continue to execute action $a$ at the motion level. The goal state is sampled from state mapping function in Line 9. Line 10-14 is a loop to execute the action. Specifically, the motion planner will compute and execute a desired control signal $\langle \delta, \theta \rangle$ repeatedly until the vehicle reaches the goal pose (Line 10). The vehicle’s current pose $x$ will be updated after each execution (Line 13). After completing the operation, the tuple $\langle s, a \rangle$ will be removed from the plan $p$ (Line 15). On the contrary, if the task planner suggests a new plan $p'$ different from the plan $p$, the currently optimal $p'$ will replace the non-optimal plan $p$ (Line 17).

C. Algorithm Instantiation

Task Planner: Our task planner $P^t$ is implemented using Answer Set Programming (ASP), which is a popular declarative language for knowledge representation and reasoning, and ASP has been used for task planning [32], [33], [11], [34]. For example, predicate $\text{leftof}(L_1, L_2)$ can be used to specify left lane $L_1$ being on the left of lane $L_2$. We model five driving actions, including $\text{mergeleft}$, $\text{mergeright}$, $\text{forward}$, $\text{turnleft}$, and $\text{turnright}$. For instance, action $\text{mergeright}$ can be used to help the vehicle merge to the right lane, where constraints, such as “$\text{changeright}$” is allowed only if there exists a lane on the right, have been modeled as well.

Motion Planner: At the motion level, path planner firstly generates a desired continuous trajectory with the minimal traveling distance using A* search. The trajectory includes a set of waypoints (each in the form of a pair of $x - y$
Using the updated cost function, the task planner re-computes an optimal plan. The vehicle plans to avoid the risky behavior of merging left in Area 1, where the safety estimator reports a low safety value based on the local road condition. This computed safety value is incorporated into task planner’s cost function, and this is where the task planner integrates the safety value into its cost function, and re-computes an optimal plan, Plan B. Different from Plan A, Plan B suggests the vehicle to go straight, and merge left in Area 2. In this trial, the vehicle was able to follow Plan B all the way to the goal. TMPUD enabled the vehicle to avoid the risky behavior of merging left in Area 1 without introducing extra motion cost. A demo video is provided on YouTube.  

A. Full and Abstract Simulation Platforms

Experiments conducted in CARLA are referred as being in full simulation. All vehicles move at a constant speed (20 km/h) on average. In full simulation, ego vehicle performs the whole plan at the task level in a the presence of other vehicles. We spawn different numbers of vehicles (200 and 120), and refer to traffic of the two environments as being heavy and normal respectively. Running full simulation using CARLA is time-consuming, preventing us from conducting large-scale experiments. For instance, results reported in this paper are based on tens of thousands of experimental trials, and full simulation in this scale would have required months of computation time. To conduct large numbers of experimental trials, we developed an abstract simulation platform, where action outcomes are sampled from pre-computed probabilistic world models. Parameters of the world models (for abstract simulation) are learned by repeatedly spawning the ego and surrounding vehicles in a small area, and statistically analyzing the results of the vehicles’ interaction.

In particular, we spent the most effort in analyzing the outcomes of “merging lane” actions due to its significant potential risks. We empirically computed the probabilities of the three different outcomes of “merging lane” actions, including “merge”, “collide”, and “stop”. We introduced two domain factors into the abstract simulation platform, including density and acceleration. In high-density environments, the ego vehicle is surrounded by three vehicles, while this number is reduced to one in low-density environments. In high-acceleration environments, surrounding vehicles’ acceleration (in m/s²) is randomly sampled in [-1.0, 1.0], while this range is [-0.5, 0.5] in low-acceleration environments.

B. Evaluation Metrics and Two Baseline Methods

The goal of TMPUD is to improve task-completion efficiency (to reduce traveling distance), while guaranteeing safety. So, the two most important evaluation metrics are traveling distance and the number of unsafe behaviors, where unsafe behaviors cause either collisions or force at least one surrounding vehicle to stop (to avoid collisions).

The two baseline methods used in this research are selected from the literature, and referred to as No-communication (No-com), and Threshold-based (Th-based). The No-com baseline [13] forces the vehicle to execute all task-level actions at the motion level, while driving behaviors’ safety values are not considered. The Th-based baseline [8] enables the motion planner to “reject” a task-level action when its safety value is lower than a threshold.

Illustrative Example: Fig. 3 presents an illustrative example. TMPUD starts with using our optimal task planner to compute Plan A. The vehicle takes the first symbolic action from Plan A (trajectory in blue color), and executes the action using our motion planner. Getting close to Area 1, the vehicle plans to merge left. However, the safety estimator at the motion level reports a low safety value in Area 1. This computed safety value is incorporated into task planner, where the task planner integrates the safety value into its cost function, and re-computes an optimal plan, Plan B.
and motion planners, and evaluate their performances in different testing platforms (e.g., using simulators with a

Fig. 4. **Abstraction simulation:** the overall performances of TMPUD and two baseline methods. The x-axis represents the average traveling distance, and the y-axis represents the total number of collisions and stops. The four subfigures correspond to four different road conditions. The road conditions, from left to right, are low-density and low-acceleration, low-density and high-acceleration, high-density and low-acceleration, high-density and high-acceleration. Under each road condition, we evaluate each algorithm using 4000 trials. We did batch-based evaluations with four batches for significance analysis, where each batch includes 1000 trials.

### Table I

<table>
<thead>
<tr>
<th>Traffic Condition</th>
<th>Algorithm</th>
<th>Traveling Distance (m)</th>
<th>Num. of collisions and stops</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal Traffic</td>
<td>TMPUD</td>
<td>514</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Th-based</td>
<td>β = 0.2</td>
<td>537</td>
</tr>
<tr>
<td></td>
<td></td>
<td>β = 0.4</td>
<td>513</td>
</tr>
<tr>
<td></td>
<td></td>
<td>β = 0.6</td>
<td>478</td>
</tr>
<tr>
<td></td>
<td>No-com</td>
<td></td>
<td>426</td>
</tr>
<tr>
<td>Heavy Traffic</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>TMPUD</td>
<td></td>
<td>530</td>
</tr>
<tr>
<td></td>
<td>Th-based</td>
<td>β = 0.5</td>
<td>545</td>
</tr>
<tr>
<td></td>
<td></td>
<td>β = 0.3</td>
<td>528</td>
</tr>
<tr>
<td></td>
<td></td>
<td>β = 0.1</td>
<td>497</td>
</tr>
<tr>
<td></td>
<td>No-com</td>
<td></td>
<td>426</td>
</tr>
</tbody>
</table>

β, where a higher (lower) β threshold makes a vehicle more conservative (aggressive). In case of an action being rejected, the task planner will compute a new plan to avoid the risky action. We develop three versions of of the Th-based baseline with different β values (0.1, 0.3, and 0.5).

C. Results

**Results from Full Simulation:** Table I presents the results in comparing TMPUD to the two baseline methods. From the table, we see that, in both road conditions, TMPUD achieved the lowest traveling distance, in comparison to those methods that produced comparable safety levels (in terms of the number of collisions and stops). For instance, under normal traffic, only the Th-based baseline with β = 0.5 was able to completely avoid collisions and stops, but it produced an average traveling distance of 537m. In comparison, TMPUD required only 514m, while completely avoided collisions and stops. Under heavy traffic, TMPUD (again) produced the best performance in safety (based on the number of collisions and stops), while requiring less traveling distance in comparison to the only baseline (Th-based with β = 0.5) that produced comparable performance in safety. The experimental trials (200 for each approach) from full simulation took eight full workdays. We aim at evaluating the performance of TMPUD under different domain factors, requiring a much larger number of trials, which motivated us to conduct experiments using the abstract simulator.

**Results from Abstract Simulation:** Fig. 4 presents the performances of TMPUD and the baseline methods in both traveling distance and the number of unsafe behaviors. The x-axis corresponds to the average traveling distance, and y-axis corresponds to the total number of collisions and stops (both are considered failure cases of driving behaviors). From the four subfigures, we see that TMPUD is the most efficient (x-axis) among those methods that produced comparable performances in safety (y-axis), except that Th-based (β = 0.5) produced slightly less unsafe behaviors (but it performed poorly in efficiency).

There are a few side observation. Not surprisingly, No-com produced the worst performance of in safety (y-axis), though its traveling distance remains the lowest. This is because, using No-com, the vehicle blindly executes task-level actions while unrealistically believing driving behaviors are always safe. The Th-based baseline’s performance depends on its safety threshold (β), where a greater value produces safer but less efficient behaviors. The results support our claim that TMPUD improves vehicles’ task-completion efficiency, while ensuring safety in different road conditions.

VI. CONCLUSIONS AND FUTURE WORK

In this paper, focusing on urban driving scenarios, we develop a safety evaluation algorithm, and a task-motion planning algorithm, called TMPUD, for autonomous driving. TMPUD, for the first time, bridges the gap between task planning and motion planning in autonomous driving. We have extensively evaluated TMPUD using a 3D urban driving simulator (CARLA) and an abstract simulator. Results suggest that TMPUD improves the task-completion efficiency in different road conditions, while ensuring the safety of driving behaviors.

In the future, we will implement TMPUD using different task and motion planners, and evaluate their performances in different testing platforms (e.g., using simulators with a
physics engine) under different conditions. Also, there is the possibility of implementing and evaluating TMPUD using indoor mobile robots.

ACKNOWLEDGEMENTS

The authors thank Fangkai Yang for discussions on this topic. This work has taken place in the Autonomous Intelligent Robotics (AIR) Group at SUNY Binghamton. AIR research is supported in part by grants from the National Science Foundation (IIS-1925044 and REU Supplement), Ford Motor Company (URP Awards), OPPO (Faculty Research Award), and SUNY Research Foundation.

REFERENCES


