2015 IEEE Intelligent Vehicles Symposium (IV)
June 28 - July 1, 2015. COEX, Seoul, Korea

Situation-aware Decision Making for Autonomous Driving on Urban Road using Online POMDP

Wei Liu Seong-Woo Kim Scott Pendleton Marcelo H. Ang Jr.

***Abstract*— As autonomous vehicles begin venturing on the urban road, rational decision making is essential for driving safety and efficiency. This paper presents a situation-aware decision making algorithm for autonomous driving on urban road. Specifically, an urban road situation model is proposed first for proper environment representation, thereafter the situation-aware decision making problem is modeled as a Partially Observable Markov Decision Process (POMDP) and solved in an online manner. The proposed algorithm has been extensively evaluated, which is general enough for autonomous driving in various urban road scenarios, including leader following, collision avoidance and traffic negotiation at both T-junction and roundabout.**

I. Introduction

As autonomous vehicles operating on the urban road, the decision making module is in charge of properly maneuvering the autonomous vehicle's own behavior and negotiating with other traffic participants. Toward this end, full awareness of the surrounding environment is necessary, but difficult to achieve however, as it is always aggravated by both the lack of proper environment representation and imperfect perception [1].

Unlike human drivers whom have the inborn ability of perceiving the environment and extracting essential information to make driving decisions, a well-designed environment model is always required by most of the state-of-art autonomous vehicles [2]. Due to the typical diversity of the environment, to design a comprehensive representation can be quite challenging. Moreover, a good environment representation should incorporate all the necessary information for proper situation evaluation - thereby to judge whether enough information has been incorporated or not is another challenge. Thirdly, since the operating environment of autonomous vehicles can be scaled up to a quite large area of coverage, the scalability of the environment representation needs to be guaranteed as well.

Another concern is incomplete perception, which can arise from various aspects. In the metric level, the vehicle perception experiences difficulties introduced by inevitable sensor noise and sensing occlusion. The resulting sensing uncertainties can dramatically weaken the autonomous vehicle's ability to perceive obstacles' metric states like pose and speed. Upgrading the vehicle perception to the semantic level, the reasoning of the other traffic participants' driving intentions is always impeded by the diversity of their behaviors and the lack of an "intention" sensor [3]. Thereby the obstacles' future behavior and the temporal evolution of the environment cannot be predicted with certainty, posing a great challenge for autonomous driving decision making.

In sight of these challenges, this paper proposes a situation-aware decision making algorithm for autonomous driving in the urban road environment. More specifically, an *urban road situatio*n model is proposed first for proper environment representation, which is defined as the integration of the road context and the obstacle vehicle's motion intention. Thereafter, the situation-aware decision making problem is modeled as a POMDP to handle the uncertainties introduced by both the motion intention and the perception noise. For the purpose of real-time application, the online POMDP solver DESPOT [4] is employed for its efficiency to account for the large observation space. The proposed algorithm has been extensively validated, and is general enough to account for various urban road scenarios, like leader following on single-lane road, collision avoidance, and traffic negotiation at the T-junction and roundabout. The contribution of this study therefore can be summarized as:

- We explicitly integrate the road context and the obstacle vehicle's motion intention into an urban road situation model, which encapsulates enough information for proper driving decision making on urban road.
- The motion intention is inferred from the vehicle *reactions*, i.e., the *deviation* of the observed vehicle states to the *reference vehicle behaviors* that are represented by the road context, where generality thereby is greatly improved.
- To model the situation-aware decision making problem as an online POMDP, the uncertainties arising from motion intention and perception noise are properly acknowledged and handled.

The remainder of this paper is organized as follows. Section II introduces the related works and the environment representation is discussed in Section III. In Section IV, a POMDP model for situation-aware decision making is presented. The evaluation results are shown in Section V and this paper is concluded in Section VI.

II. Related Work

Having been recognized as an essential component of autonomous vehicles, the autonomous driving decision making problem has been the focus of many research works in the

Wei Liu, Scott Pendleton and Marcelo H. Ang Jr. are with Department of Mechanical Engineering, National University of Singapore, Kent Ridge, Singapore {liu_wei, scott.pendleton01, mpeangh}@ nus.edu.sg

Seong-Woo Kim is with Seoul National University, Seoul, South Korea sinabrlo@ snu.ac.kr

This research was supported by the Future Urban Mobility project of the Singapore-MIT Alliance for Research and Technology (SMART) Center.

past decade. The most common approach is to manually tailor the specific action sets for different situations using a Finite State Machine or similar frameworks like behavior trees [2], [5]. As the winner of the DARPA Urban Challenge (DUC), the vehicle *BOSS* designed a state machine for behavior control, where the action transitions were triggered by events characterized by the obstacle vehicles' metric states [2]. Similar framework can be found in Junior [5], the Stanford's entry to DUC. As a milestone of autonomous vehicle development, this framework has been widely adopted by many vehicle platforms [6]. While great achievements have been made, these approaches are functioning in a reactive and greedy manner, which lacks the comprehensive understanding of the environment. The absence of full situational awareness can make the driving decisions become danger-prone, as evidenced by an incident of DUC [7]. Moreover, sensing uncertainties have been occasionally overlooked, thereby the robustness of such methods cannot be guaranteed.

Recognizing the shortcomings of reactive decision making, the research interests on obstacle behavior analysis and its extension for intelligent decision making have been increasing recently [8], [9]. POMDP and its variants have been widely adopted to account for the inevitable uncertainty of vehicle behavior prediction. An intention-aware motion planning algorithm using MOMDP is presented in [3], where the obstacle's motion intention is defined as their targeted destinations, and the obstacle's motion is assumed to be driven by the corresponding intention. As an extension, a similar work was applied to a clustered pedestrian environment in [10]. While these works achieved success in new complex applications, utilizing the targeted destinations to represent motion intention lacks generality, because the target destinations, to a certain degree, are restricted to the specific environment settings.

Besides intention-aware planning, the POMDP model has been employed for many other autonomous driving functions. Brechtel *et. al.* presented a probabilistic decision making module using continuous POMDP in [1], whose interest is on balancing the exploration and exploitation to account for the incomplete perception issue, yet the vehicle behavior uncertainty, however, is not explicitly considered. Similarly, Wei *et. al.* presented a QMDP based approach to account for the perception constraints in [11]. The work proposed in [12] aimed to employ POMDP to solve the decision making problem for lane change behavior, however the method is too specific for other situations.

As reviewed above, a general situation-aware decision making approach which can account for various autonomous driving scenarios is still an open problem.

III. Urban Road Situation

In an urban environment, an autonomous vehicle's driving decisions should be based on the road context and the other vehicles' driving behaviors. The knowledge of the road context can help to provide a more comprehensive understanding of the environment, such as the road topology, typical vehicle motion pattens and so forth. The road context thereby can be applied to either constrain the autonomous vehicle's own maneuver or narrow down the reasoning scope of the other traffic participants' behaviors. Similarly, to understand the semantic meaning of the other vehicles' behaviors, i.e., motion intention, can contribute to a more robust prediction of their future motions, which in turn can be employed for a more accurate risk evaluation. Therefore, a loose definition of the *Urban Road Situation* $\mathcal{URS}$ is proposed as *the integration of the road context* $\mathcal{C}$ *and the obstacle vehicle's motion intention* $\mathcal{I}$, namely $\mathcal{URS} = \{\mathcal{C}, \mathcal{I}, \mathcal{C} \bigoplus \mathcal{I}\}$, where the operator $\bigoplus$ indicates the correlation between the road context and the vehicle's motion intention. As such, the urban road situation features not only the motion intention and the road context, but also the implicit correlations between them.

A. Road Context

In order to properly model the urban road features, the road context should comprise not only the road network and traffic rules, but also the typical vehicle motion patterns. In our previous work in [13], the road context is inferred in a probabilistic manner using an amount of randomly observed vehicle states, where the resulting road context $\mathcal{C}$ is given in a tuple $\mathcal{C} = \{\mathbf{E}, \mathbf{R}, \mathbf{GP}\}$, where $\mathbf{E} = \{e_1, e_2, ..., e_N\}$ denotes a set of mutually exclusive regions decomposing the road surface $\mathcal{X}_R \in \mathbb{R}^2$,

$$\bigcup_{i=1}^{N} e_i = \mathcal{X}_R, e_i \cap e_j = \emptyset, \forall i \neq j \in \{1, 2, ..., N\}. \quad (1)$$

The transition probability matrix $\mathbf{R}_{N\times N}$ models the traffic rules dominated transition probability between any regions $e \in \mathbf{E}$. Formally speaking, let ρ denotes the traffic rules, the transition probability from region $e_i \in \mathbf{E}, i \in \{1, ..., N\}$ to any other region $e_j \in \mathbf{E}, \forall j \in \{1, ..., N\}$ is featured as,

$$\mathbf{R}_{i,j} = \Pr(e_j | e_i, \rho), \sum_{j=1}^{N} \mathbf{R}_{i,j} = 1, \forall i, j \in \{1, 2, ..., N\}. \quad (2)$$

Moreover, $\mathbf{GP}$ represents the *Gaussian Process Regression* model that generalizes a time-independent representation of the vehicle speed distribution over the entire road surface $\mathcal{X}_R$. Specifically, let $p = [\mathrm{x}, \mathrm{y}] \in \mathcal{X}_R$ denote the vehicle position vector, the corresponding posterior distribution of the generalized speed v_{gp} is mapped from $\mathbf{GP}$ as,

$$\mathrm{v}_{\mathrm{gp}}(p) \sim \mathbf{GP}(p) = \mathcal{N}(\mathbf{m}(p), \mathbf{K}(p)), \quad (3)$$

where the $\mathrm{v}_{\mathrm{gp}}(p)$ is distributed as a Gaussian with mean function $\mathbf{m}(p)$ and covariance function $\mathbf{K}(p)$. The readers can refer to [13] for more details.

Given the observed vehicle state s_o, the corresponding road context information $c(s_o) = \{e(s_o), \mathbf{R}(s_o), \mathrm{v}_{\mathrm{ref}}(s_o)\}$ can be mapped out as,

$$\begin{aligned} e(s_o) &= \{e \in \mathbf{E} | s_o.p \in e\}, \\ \mathbf{R}(s_o) &= \mathbf{R}_{e(s_o), \forall e' \in \mathbf{E}}, \\ \mathrm{v}_{\mathrm{ref}}(s_o) &\sim \mathbf{GP}(s_o.p), \end{aligned} \quad (4)$$

where $s_o.p \in \mathcal{X}_R$ denotes the vehicle position vector and $e(s_o)$ is the region where the vehicle locates. The

$\mathbf{R}_{e(s_o),\forall e' \in \mathbf{E}}$ denotes the transition probability vector that corresponds to the region $e(s_o)$, and the speed $\mathrm{v}_{\mathrm{ref}}(s_o)$ is generated from the Gaussian Process Regression model $\mathbf{GP}$. The road context $\mathcal{C}$ essentially is a generalization of the vehicles' behaviors derived by abstracting the vehicle behavior consistencies, thus it can properly represent the normal vehicle behavior that should be followed. Specifically, the road context information $c(s_o)$ not only regulates the directions toward which the observed vehicle is allowed to move, but also provides the reference speed $\mathrm{v}_{\mathrm{ref}}$ that the observed vehicle should follow. Therefore, the $c(s_o)$ can be defined as the *reference vehicle behavior* that corresponds to the observed vehicle state s_o.

B. Motion Intention

The most popular strategy of vehicle motion intention modeling is to define the motion intention as their hidden destinations in Fig. 1(a). Namely, the obstacle vehicle's final destination is not observable and there always exists a motion model $m \in \mathbf{M}$ that dominates the vehicle's future motion for each hidden goal $g \in \mathbf{G}$, where $\mathbf{M}$ and $\mathbf{G}$ denotes the motion model set and hidden goal set respectively. More specifically, let $\mathcal{ROI}s$ be the *Regions Of Interest* where the autonomous vehicle is primarily concerned about the obstacle vehicles' future motion. Without losing generality, assume that the autonomous vehicle is operating in an environment that includes M $\mathcal{ROI}s$, which can include T-junctions, roundabouts and so forth. For each $\mathcal{ROI}^{[i]}, i \in \{1, ..., M\}$, there exists a goal set $\mathbf{G}^{[i]} \leftarrow \mathcal{ROI}^{[i]}, i \in \{1, ..., M\}$ that defines the possible directions where the vehicles can move. Therefore, the size of the required motion models can be given as $|\mathbf{M}| = \sum_{i=1}^{M} |\mathbf{G}^{[i]}|$, where $|\mathbf{G}^{[i]}|$ denotes the size of the goal set $\mathbf{G}^{[i]}, i \in \{1, ..., M\}$. Given $M \rightarrow \infty$, the size of $\mathbf{M}$ becomes unbounded, which implies that using the hidden goal to model the motion intention may suffer from the scalability problem as the amount of $\mathcal{ROI}s$ increases.

This study aims at employing the obstacle vehicle's *reaction* to model the motion intention as Fig. 1(b). This proposal is inspired by the fact that the vehicle behaviors are aways correlated, such that human drivers reactively adjust their driving strategies when interacting with each other. The most intuitive method of reaction modeling might be measuring the temporal evolutions of the vehicle state, such as the route changing, speed variance, etc.. To properly quantify these changes, however, is complicated due to the diversity of vehicle behaviors. In this study, the *reaction* φ is proposed as the *deviation* of the observed vehicle state s_o to the corresponding reference vehicle behavior $c(s_o)$, i.e., $\varphi \leftarrow s_o \bigotimes c(s_o)$, where the operator $\bigotimes$ measures the deviations. The rationality of this proposal can be evidenced by the fact that the reactive driving strategies, in most of cases, are inconsistent with the normal driving behaviors captured by the road context.

Given the deviation measurement, the reaction based motion intention is proposed as $\mathcal{I} \sim \Pr(\mathcal{I}|s_o \bigotimes c(s_o))$. As such, the motion intention is abstracted as the conditional distribution over the reactions, where the implicit correlations

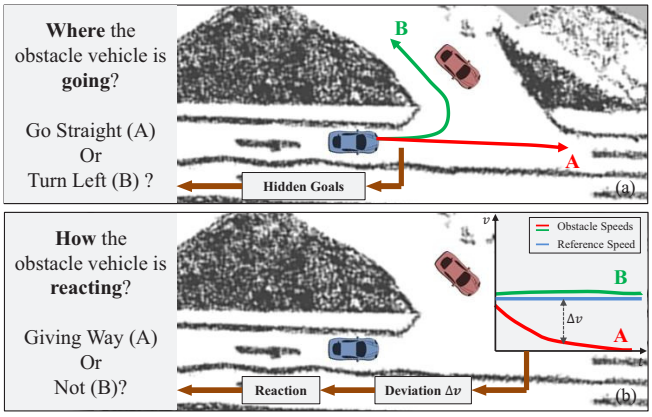

Fig. 1. Motion intention explanation: (a) Hidden goal based motion intention, where A and B denotes two optional routes for the obstacle vehicle (blue). (b) Reaction based motion intention (using speed deviation), where A and B denotes two possible obstacle vehicle speed profiles, which are checked against the reference speed for motion intention inference.

between the road context and motion intention are acknowledged. Compared to the hidden goal method, employing the reactions to model the motion intention is more general and scalable, because the measurement of deviation is not restricted to any specific $\mathcal{ROI}$. Thereby we can always find a bounded-size reaction based motion intention set $\mathcal{I}$ to model the vehicle behaviors.

In our study, the vehicle speed deviation is utilized and the motion intention set with four hypothesis is designed as,

$$\mathcal{I} = \{Stopping, Hesitating, Normal, Aggressive\},$$

where their meanings and the corresponding speed deviations are defined as follows:

- The *Stopping* intention indicates that the obstacle vehicle plans to commit a stop for temporal parking or giving way to approaching vehicle. As a feature, the obstacle vehicle's speed is close to zero, while the reference speed is much higher.
- The *Hesitating* intention indicates that the obstacle vehicle is hesitating on making driving decisions. The speed deviation is given as that the obstacle vehicle speed is slower than the reference speed but not close to zero. This is inspired by human driving activity, where slower speeds are usually preferable when the current driving decisions are not confident.
- The *Normal* intention indicates that the obstacle vehicle will maintain its current normal behavior. As an evidence, the obstacle vehicle's speed is matching well with the reference speed, namely the speed deviation is small.
- The *Aggressive* intention indicates that the obstacle vehicle may aggressively accelerate and has no sense for negotiation. As an indicator, the obstacle vehicle's current speed is much higher than the reference speed.

While these motion intention definitions may be questionable the first sight, they are abstracted from human driving behaviors and the functionality has been validated by our extensive evaluations.

IV. Situation-aware Decision Making

Recognizing the considerable uncertainties arising from the urban road environment, stochastic planning strategies are necessary for robust driving decision making. As a principled general framework for acting and planning in a partially observable environment, the POMDP is adopted in this study. This equips the autonomous vehicle with the ability to estimate and evaluate the outcome of driving decisions, even when the situation cannot be exactly observed.

A. Preliminary on POMDP

A POMDP is formally a tuple $\{\mathcal{S}, \mathcal{A}, \mathcal{Z}, T, O, R\}$, where $\mathcal{S}$ is the state space, $\mathcal{A}$ is a set of actions and $\mathcal{Z}$ denotes the observation space. The transition function $T(s', s, a) = \Pr(s'|s, a) : \mathcal{S} \times \mathcal{A} \times \mathcal{S}$ models the probability of transiting to state $s' \in \mathcal{S}$ when the agent takes an action $a \in \mathcal{A}$ at state $s \in \mathcal{S}$. The observation function $O(z, s', a) = \Pr(z|s', a) : \mathcal{S} \times \mathcal{A} \times \mathcal{Z}$, similarly, gives the probability of observing $z \in \mathcal{Z}$ when action $a \in \mathcal{A}$ is applied and the resulting state is $s' \in \mathcal{S}$. The reward function $R(s, a) : \mathcal{S} \times \mathcal{A}$ is the reward obtained by performing action $a \in \mathcal{A}$ in state $s \in \mathcal{S}$.

The solution to the POMDP thereby is an optimal policy π^* that maximize the expected accumulated reward $E(\sum_{t=0}^{\infty} \gamma^t R(a_t, s_t))$, where $\gamma \in [0, 1)$ is a discount factor, s_t and a_t denote the agent's state and action at time t. The true state, however, is not fully observable to the agent, thus the agent maintains a belief state $b \in \mathcal{B}$, i.e. a probability distribution over $\mathcal{S}$, instead. The policy $\pi : \mathcal{B} \rightarrow \mathcal{A}$ therefore maps a prescribed action $a \in \mathcal{A}$ from a belief $b \in \mathcal{B}$.

B. A POMDP Model for Situation-aware Decision Making

This section details how the situation-aware decision making problem can be fitted and solved as a POMDP by elaborating on each component of the contextualized POMDP tuple.

1) State Space $\mathcal{S}$: Due to the Markov property, the state space $\mathcal{S}$ must hold sufficient information for either decision making or belief update [1], which in this case comprises the vehicle pose $[\mathrm{x}, \mathrm{y}, \theta] \in \mathbb{R}^2 \times \mathbb{S}$ and vehicle speed $\mathrm{v} \in \mathbb{R}$ for all the vehicles involved. For the obstacle vehicles, the motion intention $\iota \in \mathcal{I}$ also needs to be covered by the obstacle vehicle state for proper state transition modeling. The road context $\mathcal{C}$, on the other hand, can be employed as a static reference knowledge, and is consequently excluded from the vehicle state.

Formally speaking, the joint state $s \in \mathcal{S}$ can be given as,

$$s = [s_e, s_1, s_2, ..., s_K]^{\mathrm{T}}, \tag{5}$$

where s is composed by the ego vehicle state s_e and the obstacle vehicles' states $s_i, \forall i \in \{1, 2, ..., K\}$, and K is the number of the obstacle vehicles involved. Let the vehicle metric state $x \in \mathbb{R}^3 \times \mathbb{S}$ be defined as $x = [\mathrm{x}, \mathrm{y}, \theta, \mathrm{v}]^{\mathrm{T}}$, which includes the vehicle pose and speed. The ego vehicle state can be equivalently defined as $s_e = x_e$. Similarly, the obstacle vehicle state s_i is given as $s_i = [x_i, \iota_i]^{\mathrm{T}}$, where the vehicle motion intention ι_i is explicitly modeled. Given the recent research advances in vehicle detection and tracking

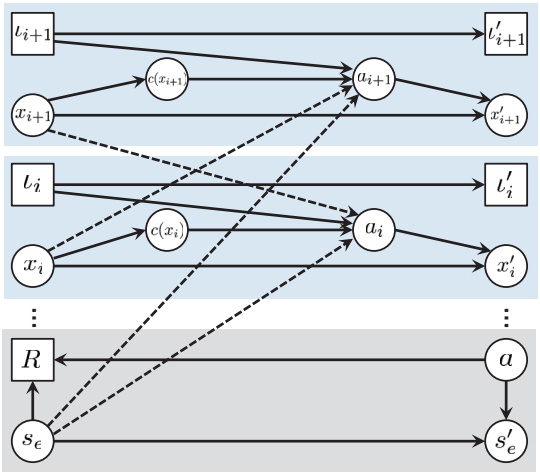

Fig. 2. Transition model for situation-aware decision making. The dashed lines are to model the vehicle correlations.

[14], we assume that the vehicle metric state x can be properly observed with Gaussian noise imposed, however the motion intention is only partially observable.

2) Action Space $\mathcal{A}$: In our autonomous vehicle navigation system, the decision making module responds for planning some tactical maneuvers, such as traffic negotiation at a T-junction or roundabout, by maintaining the safe distance with other obstacle vehicles while following a reference route. Therefore the action space $\mathcal{A}$ can be covered by a discrete action set consisting of acceleration, deceleration or maintaining current speed as,

$$\mathcal{A} = [Acc., Dec., Cur.]^{\mathrm{T}}. \tag{6}$$

As such, the defined actions are in charge of the vehicle speed control. The steering control of the ego vehicle, on the other hand, can be accomplished by tracking the reference routes closely.

3) Observation Space $\mathcal{Z}$: Similar to the joint state $s \in \mathcal{S}$, the joint observation $z \in \mathcal{Z}$ consists of the following elements,

$$z = [z_e, z_1, z_2, ..., z_K]^{\mathrm{T}}, \tag{7}$$

where z_e is the ego vehicle's observation and z_i denotes the observation of obstacle vehicle i. To properly update the obstacle vehicle's motion intention belief, each vehicle's observation consists of its pose and speed. Since the vehicle metric state x can be properly observed, we can simply generate the observations with an one-on-one mapping directly from the corresponding metric states.

4) Transition Model $T(s, a, s')$: The transition function describes the stochastic system dynamics driven by both the action applied to the ego vehicle and the obstacle vehicles' motion intention. The overall structure of the transition model is illustrated by the Bayesian Network in Fig. 2.

Formally speaking, the transition model can be represented by the probabilistic transition as,

$$\Pr(s'|s, a) = \Pr(s'_e|s_e, a) \prod_{i=1}^{K} \Pr(s'_i|s), \tag{8}$$

where $a \in \mathcal{A}$ is the action applied to the ego vehicle and s' is the new joint state. For ego vehicle, the state transition $\Pr(s'_e|s_e, a)$ is dominated by the applied action a and the turning angle $\Delta\theta$ only, which can be given in detail as,

$$\begin{bmatrix} \mathrm{x}'_e \\ \mathrm{y}'_e \\ \theta'_e \\ \mathrm{v}'_e \end{bmatrix} = \begin{bmatrix} \mathrm{x}_e \\ \mathrm{y}_e \\ \theta_e \\ \mathrm{v}_e \end{bmatrix} + \begin{bmatrix} (\mathrm{v}_e + a\Delta t)\Delta t \cos(\theta + \Delta\theta) \\ (\mathrm{v}_e + a\Delta t)\Delta t \sin(\theta + \Delta\theta) \\ \Delta\theta \\ a\Delta t \end{bmatrix}, \quad (9)$$

where $\Delta\theta$ is achieved by tracking the reference route.

For the obstacle vehicles, the state transition model $\Pr(s'_i|s)$ is more complicated as the obstacle vehicle's motion intention is not observable and the dependencies within the vehicles need to be carefully accounted for. Recalling our discussion in Section III, the vehicles' behavior is driven by both the road context and their motion intentions, therefore an obstacle vehicle's state transition can be factorized as,

$$\begin{aligned} \Pr(s'_i|s) &= \Pr(x'_i|s)\Pr(\iota'_i|\iota_i) \\ &= \sum_{a_i} \Pr(x'_i|x_i, a_i)\Pr(a_i|s)\Pr(\iota'_i|\iota_i), \end{aligned} \quad (10)$$

where a_i is the action applied to obstacle vehicle i. Unlike the ego vehicle action $a \in \mathcal{A}$, the obstacle vehicle action a_i is defined as $a_i = [a_{vi}, \Delta\theta_i]^{\mathrm{T}} \in [Acc., Dec., Cur.] \times \mathbb{S}$, which includes both the speed action a_{vi} and turning angle $\Delta\theta_i$. Given the obstacle vehicle action a_i, the obstacle vehicle's metric state transition $\Pr(x'_i|x_i, a_i)$ is identical to that of the ego vehicle in Eqn. (9). Regarding the motion intention transition $\Pr(\iota'_i|\iota_i)$, the motion intention remains unchanged during the state transition process, which will be updated accordingly when a new observation is available.

The difficulty then lies on the inference of the obstacle vehicle action a_i. Let $\bar{X}_i = \bigcup_{j \neq i} x_j, \forall j \in \{e, 1, ..., K\}$ represents all the vehicles' metric states excluding the one of vehicle i, thereafter the action distribution $\Pr(a_i|s)$ in Eqn. (10) can be reformulated as,

$$\begin{aligned} \Pr(a_i|s) &= \Pr(a_{vi}, \Delta\theta_i|\bar{X}_i, x_i, \iota_i) \\ &= \sum_{c(x_i)} \Pr(a_{vi}, \Delta\theta_i|\bar{X}_i, x_i, \iota_i, c(x_i))\Pr(c(x_i)|x_i) \\ &= \sum_{c(x_i)} \Pr(\Delta\theta_i|c(x_i))\Pr(a_{vi}|\bar{X}_i, x_i, \iota_i)\Pr(c(x_i)|x_i), \end{aligned} \quad (11)$$

where $c(x_i) = \{e(x_i), \mathbf{R}(x_i), \mathrm{v}_{\mathrm{ref}}(x_i)\}$ is the reference vehicle behavior sampled from Eqn. (4). Since the obstacle vehicle's moving direction is not explicitly modeled by the motion intention, its turning angle $\Delta\theta_i$ thereby is controlled by the road context $c(x_i)$ only as $\Pr(\Delta\theta_i|c(x_i))$. More specifically, given the current region $e(x_i)$ inside which the obstacle vehicle i locates, we can sample the next possible region $e(x_i)' \in \mathbf{E}$ where the obstacle vehicle might locate according to $\mathbf{R}(e(x_i)) = \Pr(e(x_i)'|e(x_i), \rho)$, thereafter the turning angle is calculated as $\Delta\theta_i = \mathcal{H}(e(x_i), e(x_i)') - \theta_i$, where the function $\mathcal{H}(e(x_i), e(x_i)')$ calculates the orientation angle between the region $e(x_i)$ and $e(x_i)'$.

Note that the speed action a_{vi} depends on both the motion intention and the correlations with the other vehicles. Toward this end, the dependency within the vehicles is arguably modeled as *each vehicle has to avoid the potential collision with any other vehicles*. Therefore, the speed action a_{vi} is modeled as $\Pr(a_{vi}|\iota_i, \bar{X}_i, x_i) \sim \Pr(a_{vi}|\iota_i)[\Pr(a_{vi} = Dec.|\bar{X}_i, x_i)]^{\tau}$, where $\tau = 1$ iff the distance between the obstacle vehicle i and its nearest vehicle is less than certain safety margin, otherwise $\tau = 0$. As such, if the obstacle vehicle i has a extremely low collision risk with other vehicles, the speed action is mapped from the motion intention only. In accordance with the motion intention definitions in Section III-B, the velocity action corresponding to each motion intention is given in Table I. Once the collision risk is high enough, the obstacle vehicle would either follow the inferred motion intention or choose the *Decelerate* action. The probability of selecting $Decelerate$ action is given as $\Pr(a_{vi} = Dec.|\bar{X}_i, x_i) = \xi \mathrm{v}_{nearest}/\mathrm{v}_{max}$, where $\xi \in (0, 1]$ is the scaling factor. The $\mathrm{v}_{nearest}$ denotes the nearest vehicle's speed and v_{max} is the maximum speed allowed. That is to say, the higher speed of the nearest vehicle, the higher probability of vehicle i would decide to decelerate.

TABLE I
ACTION AND MEAN FUNCTION W.R.T. MOTION INTENTION

$\iota \in \mathcal{I}$	*Stopping*	*Hesitating*	*Normal*	*Aggressive*
a_v	*Dec.*	*Dec./Acc.*[a]	*Cur.*	*Acc./Cur.*[a]
$\mathbf{m}(\mathrm{v}_{\mathrm{ref}}, \iota)$	0.0	$0.5\mathrm{v}_{\mathrm{ref}}$	$\mathrm{v}_{\mathrm{ref}}$	$1.5\mathrm{v}_{\mathrm{ref}}$

a: Each action has 50% probability to be selected.

5) Observation Model $O(z, s', a)$: The observation model aims at simulating the measurement process, which can be employed to update the motion intention belief. For each vehicle, the observation is conditionally independent with other variables given the corresponding vehicle state, thereby the observation model can be factorized as,

$$\Pr(z|s', a) = \Pr(z_e|s'_e) \prod_{i=1}^{K} \Pr(z_i|s'_i). \quad (12)$$

For the ego vehicle, the observation function is given as,

$$\Pr(z_e|s'_e) \sim \mathcal{N}(z_e|x'_e, \Sigma_{z_e}), \quad (13)$$

where the Gaussian noise is imposed to model the sensor property. The obstacle vehicles' observation functions are more complicated due to the inclusion of their motion intentions, and are thereby reformulated as,

$$\Pr(z_i|s'_i) = \Pr(z_i|x'_i, \iota'_i) = \Pr(z_i|\iota'_i)\Pr(z_i|x'_i)/\Pr(z_i), \quad (14)$$

where the metric state observation function $\Pr(z_i|x'_i)$ follows the same distribution as Eqn. (13). The distribution $\Pr(z_i|\iota'_i)$ models the probability of observing z_i given the motion intention is ι'_i, and can be given as,

$$\Pr(z_i|\iota'_i) \sim \mathcal{N}(z_{vi}|\mathbf{m}(\mathrm{v}_{\mathrm{ref}}(z_i), \iota'_i), \mathrm{v}_{\mathrm{ref}}(z_i)/\sigma)), \quad (15)$$

where z_{vi} is the observed speed of vehicle i and $\mathrm{v}_{\mathrm{ref}}(z_i)$ denotes the reference speed that corresponds to observation z_i. In other words, given the reference speed $\mathrm{v}_{\mathrm{ref}}(z_i)$, the observation function for each motion intention $\iota \in \mathcal{I}$ is modeled as a Gaussian $\mathcal{N}(z_{vi}|\mathbf{m}(\mathrm{v}_{\mathrm{ref}}(z_i), \iota), \mathrm{v}_{\mathrm{ref}}(z_i)/\sigma))$ with mean $\mathbf{m}(\mathrm{v}_{\mathrm{ref}}(z_i), \iota)$ and covariance $\mathrm{v}_{\mathrm{ref}}(z_i)/\sigma$. The

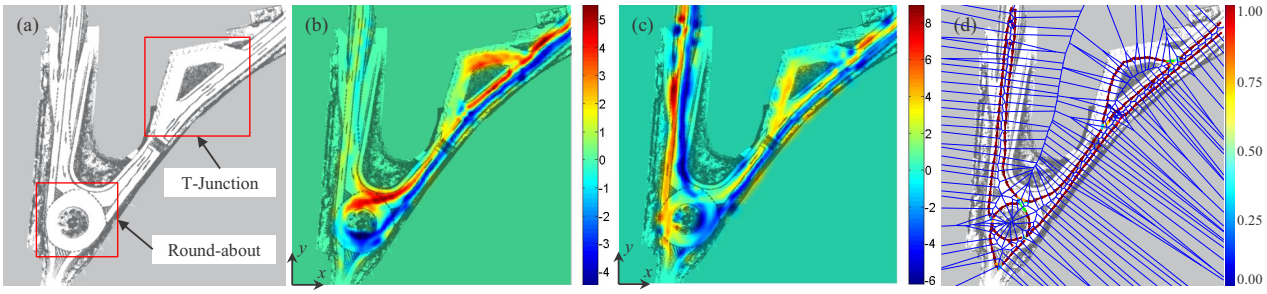

Fig. 3. Road context information: (a) shows the Occupancy Grid Map of the environment; (b) and (c) shown the mean value of the reference speed distribution in x and y direction respectively; (d) represents the region discretization and the transition probabilities indicated by the colorful edges.

mean value $\mathbf{m}(\mathrm{v}_{\mathrm{ref}}(z_i), \iota)$ is a function of the reference speed and the motion intention type as Table I, which explicitly models the relationship between the motion intention and the speed deviation. The scaling factor σ in the covariance function is used to adjust the motion intention confidence. Together with the transition function in Eqn. (10), the update of obstacle vehicle's motion intention belief thereby becomes straightforward.

6) Reward Function R: The main objective of the ego vehicle is to arrive at the target destination as quickly as possible, meanwhile avoiding collision with the other obstacle vehicles. The reward function $R(s, a)$ is thus defined in a multi-objective manner to properly balance the driving efficiency and safety,

$$R(s,a) = R_{goal}(s,a) + R_{crash}(s,a) + R_{action}(a) + R_{vel}(s),$$

where $R_{goal}(s, a)$ denotes the reward when the ego vehicle reaches the final destination and $R_{crash}(s, a)$ assigns a high penalty if the ego vehicle is in collision. $R_{action}(a)$ provides a small penalty if the $a \in [Acc., Dec.]^{\mathrm{T}}$, which is to avoid the frequent speed change and to improve comfort. The speed reward $R_{vel}(s) = k\mathrm{v}/\mathrm{v}_{max}$ is to encourage high speed travel and improve the driving efficiency, where v is the ego vehicle's speed and k is a scaling factor.

C. Online POMDP Solver

Given the POMDP model for situation-aware decision making as described in the previous sections, an efficient POMDP solver must then be applied for real-time applications.

The general POMDP discussed in Section IV-A is usually solved in an offline manner, where the optimal policy is computed offline and the agent conducts the policy looking-up for online execution. Since the algorithm aims to return a policy that prescribes an action for every possible belief state, this strategy has the risk of becoming intractable when the POMDP model is scaled up. In contrast, the online POMDP algorithm searches for a good policy given the agent's current belief state $b_0 \in \mathcal{B}$, then considers only a small set of belief states that are achievable from that current belief state, thereby greatly improving the tractability. In short, the online POMDP algorithm divides each control loop into two phases: *planning* and *execution*. The algorithm uses the current belief b_0 to compute the best action to execute in the planning phase, the action is then executed by the agent during the execution phase. Thereafter, the agent's belief state is updated according to observation obtained. This process keeps repeating until the terminating condition is satisfied. The readers can refer to [15] for a more detailed introduction of online POMDPs.

In our application, the online POMDP solver DESPOT is employed for its efficiency in handling a large observation space. Rather than searching the whole belief tree, DESPOT only samples a *scenario* set with constant size Q. As a consequence, the belief tree of height H contains only $O(|\mathcal{A}|^H Q)$ nodes, which is not correlated to the size of the observation space. This can greatly alleviate our observation space constraints. The belief state is represented by the random particles within DESPOT, and for each particle the obstacle vehicles' motion intentions are randomly sampled for a belief representation. Implementation-wise, since the belief state is represented by the particles, it is straightforward to implement both the state transition in Eqn. (8), (10) and the observation update in Eqn. (13), (14).

V. Evaluation

In this section, the proposed algorithm will be evaluated to validate its functionality and generality.

A. Settings

These evaluations are carried out on urban roads within the National University of Singapore as shown in Fig. 3(a), which includes the typical T-junction and roundabout. The corresponding road context information is shown in Fig. 3(b)-(d), where Fig. 3(b) and Fig. 3(c) represent the mean values of the generalized speed distribution in x and y direction respectively. The region discretization $\mathbf{E}$ and corresponding transition probabilities $\mathbf{R}$ are demonstrated in Fig. 3(d). These road context results are inferred from 4,600 randomly observed vehicle states using our previous work in [13]. Given the environment and the corresponding road context, the *Stage* simulator [16] is employed for evaluation, where Gaussian noise is purposely imposed on the vehicle pose and speed measurement. The Pure-Pursuit tracking algorithm is employed for the vehicle steering control [17].

B. Results

The proposed situation-aware decision making algorithm has been successfully applied for various urban road driving

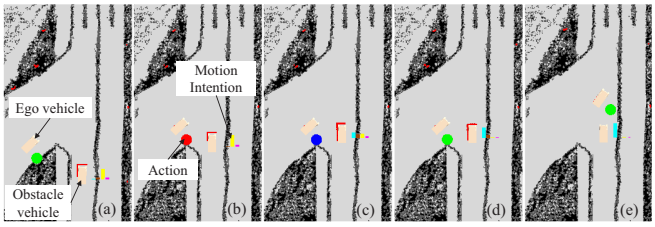

Fig. 4. Negotiation with a single obstacle vehicle which purposely gives way at the T-junction. The action is represented as the circle marker: [Red:$Dec.$, Green:$Acc.$, Blue:$Cur.$]. The motion intention is represented as the cubic marker:[Sky Blue:$Stopping$, Brown:$Hesitating$, Yellow:$Normal$, Purple:$Aggressive$], where the cubic height is proportional to the corresponding belief value.

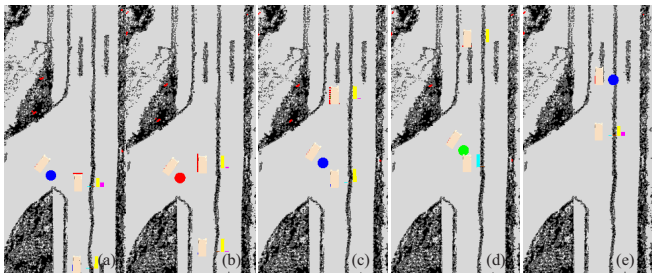

Fig. 5. Negotiation with multiple obstacle vehicles at the T-junction.

scenarios, such as leader following and negotiation with obstacle vehicles at a T-junction or roundabout. For the sake of generality evaluation, the designed POMDP model, especially the reward function, remained unchanged for all the evaluations. The demonstration of the complete evaluation results together with the evaluation parameters is available in the video at `http://youtu.be/W37haHhfU34` . In this context, we want to highlight the ability of our algorithm to handle the decision making problem at the T-junction and roundabout, which is well recognized as challenging.

As the autonomous vehicle was conducting the lane merging behavior at the T-junction shown in Fig. 4(a), it was essential to properly detect the road context and reason the obstacle vehicle's motion intention to guarantee driving safety. As illustrated in Fig. 4(b), the *Decelerate* action was triggered first to avoid the potential collision with the approaching vehicle, which was associated with a high belief of maintaining its current $Normal$ behavior. After some time, the obstacle vehicle, however, changed its mind and decided to give way to the ego vehicle as in Fig. 4(c), where its motion intention belief was updated accordingly. Given the increasing confidence that the obstacle vehicle may want to wait as Fig. 4(d), the ego vehicle decided to accelerate and cautiously merged into the desired lane as Fig. 4(e). The purpose of this scenario is to evaluate the ability of our proposed algorithm to handle complicated interactions with the other vehicles. Rather than waiting forever, the ego vehicle is able to properly reason the obstacle vehicle's intention and react in an efficient manner.

The extension of the proposed algorithm to handle multiple obstacle vehicles is shown in Fig. 5. As expected, the ego vehicle decided to decelerate and wait when the leading obstacle vehicle chose not to give way in Fig. 5(b). After the

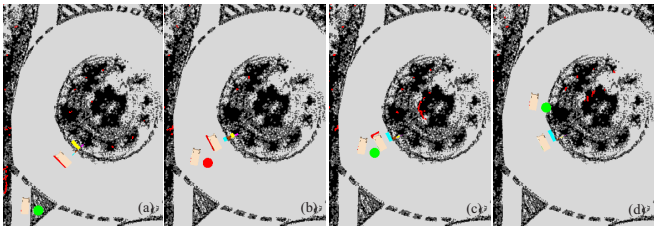

Fig. 6. Negotiation with a single obstacle vehicle which purposely gives way at the roundabout.

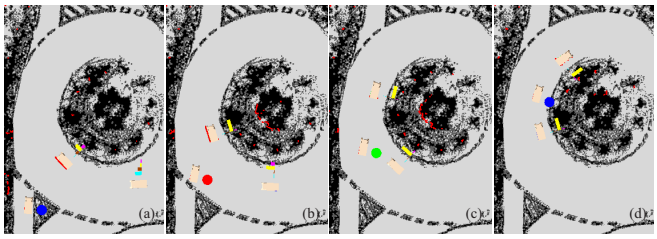

Fig. 7. Negotiation with multiple obstacle vehicles at the roundabout.

leading obstacle vehicle passing by, the ego vehicle decided to slowly move forward as Fig. 5(c), although the following obstacle vehicle was still holding a high $Normal$ belief. Eventually, the following obstacle vehicle committed a stop in order to avoid the collision with the ego vehicle, where its motion intention belief is updated as Fig. 5(d), thereafter the ego vehicle successfully merged into the gap between the two obstacle vehicles as Fig. 5(e). While actively cutting into the following obstacle vehicle's route seems dangerous, the ego vehicle's decision is within our expectation. This is because the dependency within the vehicles is well acknowledged by the ego vehicle, namely the ego vehicle knows that there would be a higher chance to make the obstacle vehicles give way if it actively cuts into their route and moves in a relatively higher speed. This is exactly the same way how the human drivers negotiate with each other.

In sight of the improvements achieved, the same model was applied for the roundabout navigation. The evaluation snapshots are shown in Fig. 6 and Fig.7, where the scenarios are identical to that of the T-junction evaluation. The similar results are achieved, which indicates that the proposed approach can function properly at the roundabout as well.

Acknowledging that the proposed approach is probabilistic in nature, we extensively conducted 100 evaluation trials for each scenario discussed above. For the single obstacle vehicle case, one additional evaluation scenario in which the obstacle vehicle will not purposely give way is implemented. Moreover, we extended the multiple obstacle vehicle case by trying different gaps between the obstacle vehicles. Two evaluation metrics including the *average traveling time before the goal is reached* and the *overall failure rate* were measured to validate the efficiency and safety respectively. The failure is defined as either a collision or failure to complete the navigation task within the permitted time. The traveling time was only measured for the successful trials. For the sake of performance comparison, we also implemented the reactive decision making method in [17], where the ego vehicle

TABLE II

NAVIGATION PERFORMANCE OF SITUATION-AWARE DECISION MAKING

Metric	Algorithm	T-Junction				Round-About			
		Single Obstacle Vehicle		Multiple Obstacle Vehicles		Single Obstacle Vehicle		Multiple Obstacle Vehicles	
		Give-way	Not Give-way	Small Gap[b]	Large Gap[b]	Give-way	Not Give-way	Small Gap[b]	Large Gap[b]
F.R.[a]	Reactive	0.18	0.00	0.02	0.00	0.15	0.00	0.00	0.01
	POMDP	0.00	0.02	0.01	0.03	0.00	0.03	0.00	0.02
T.T. [a] (sec)	Reactive	25.2	26.9	32.1	34.9	36.5	38.8	42.0	44.5
	POMDP	23.5	25.2	30.6	28.4	35.1	38.0	40.8	38.7

a: F.R.: Failure Rate; T.T.: Traveling Time. b: Small Gap: $2m \sim 3m$; Large Gap: $5m \sim 6m$.

iteratively checks the safe region for obstacle clearance measurement and reactively decides whether to go or not.

As demonstrated in Table II, the reactive approach has a quite high failure rate to handle the scenario in which the obstacle vehicle purposely gives way. This is because when the obstacle vehicle decides to give way, the distance between itself and the ego vehicle is smaller than the threshold designed for safe driving, thereby the ego vehicle just waits forever to bypass the stopped obstacle vehicle. On the other hand, our POMDP based approach can maintain an impressive success rate for all the scenarios, especially the one in which the obstacle vehicles purposely give way. Rather than getting stuck, the task failures are because of the collisions contributed by the proactive actions. As discussed earlier, these proactive actions are rational and can be well balanced by tuning the model parameters.

While the reactive approach seems able to maintain a nice success rate, the efficiency is scarified. For most of the trials using the reactive approach, the ego vehicle always passively decided to wait, even when the gap between itself and the obstacle vehicles is large enough for safe merging. This can be evidenced by the average traveling time, where the reactive approach consumes much longer time to navigate though the designed scenarios, especially the one with multiple obstacle vehicles and a large separation distance. On the other hand, the POMDP based approach can efficiently merge into the gaps between the multiple obstacle vehicles by adopting some proactive actions.

Given these consistent results, we can safely conclude that our proposed algorithm is general enough and can properly trade-off between the safety and efficiency without any manually specified rules.

VI. CONCLUSION

In this work, a situation-aware decision making approach was proposed for autonomous driving on the urban road. The concept of the urban road situation was discussed first, thereafter the situation-aware decision making problem was solved as a POMDP to improve the robustness and intelligence. For real-time application, the DESPOT algorithm was employed as the POMDP solver. The proposed approach has been extensively evaluated and its functionality and generality have been properly validated.

REFERENCES

[1] S. Brechtel, T. Gindele, and R. Dillmann, "Probabilistic decision-making under uncertainty for autonomous driving using continuous pomdps," in *Intelligent Transportation Systems (ITSC), 2014 IEEE 17th International Conference on*. IEEE, 2014, pp. 392–399.

[2] C. Urmson, J. Anhalt, D. Bagnell, C. Baker, R. Bittner, M. Clark, J. Dolan, D. Duggins, T. Galatali, C. Geyer, *et al.*, "Autonomous driving in urban environments: Boss and the urban challenge," *Journal of Field Robotics*, vol. 25, no. 8, pp. 425–466, 2008.

[3] T. Bandyopadhyay, C. Z. Jie, D. Hsu, M. H. Ang Jr, D. Rus, and E. Frazzoli, "Intention-aware pedestrian avoidance," in *Experimental Robotics*. Springer International Publishing, 2013, pp. 963–977.

[4] A. Somani, N. Ye, D. Hsu, and W. S. Lee, "Despot: Online pomdp planning with regularization," in *Advances In Neural Information Processing Systems*, 2013, pp. 1772–1780.

[5] M. Montemerlo, J. Becker, S. Bhat, H. Dahlkamp, D. Dolgov, S. Ettinger, D. Haehnel, T. Hilden, G. Hoffmann, B. Huhnke, *et al.*, "Junior: The stanford entry in the urban challenge," *Journal of field Robotics*, vol. 25, no. 9, pp. 569–597, 2008.

[6] C. Berger and B. Rumpe, "Autonomous driving-5 years after the urban challenge: The anticipatory vehicle as a cyber-physical system." *GI-Jahrestagung*, pp. 789–798, 2012.

[7] L. Fletcher, S. Teller, E. Olson, D. Moore, Y. Kuwata, J. How, J. Leonard, I. Miller, M. Campbell, D. Huttenlocher, *et al.*, "The mit–cornell collision and why it happened," *Journal of Field Robotics*, vol. 25, no. 10, pp. 775–807, 2008.

[8] S. Lefèvre, D. Vasquez, and C. Laugier, "A survey on motion prediction and risk assessment for intelligent vehicles," *Robomech Journal*, vol. 1, no. 1, pp. 1–14, 2014.

[9] W. Liu, S.-W. Kim, K. Marczuk, and M. H. Ang, "Vehicle motion intention reasoning using cooperative perception on urban road," in *Intelligent Transportation Systems (ITSC), 2014 IEEE 17th International Conference on*. IEEE, 2014, pp. 424–430.

[10] C. Shaojun, "Online pomdp planning for vehicle navigation in densely populated area," 2014.

[11] J. Wei, J. M. Dolan, J. M. Snider, and B. Litkouhi, "A point-based mdp for robust single-lane autonomous driving behavior under uncertainties," in *Robotics and Automation (ICRA), 2011 IEEE International Conference on*. IEEE, 2011, pp. 2586–2592.

[12] S. Ulbrich and M. Maurer, "Probabilistic online pomdp decision making for lane changes in fully automated driving," in *Intelligent Transportation Systems*, 2013, pp. 2063–2070.

[13] W. Liu, S.-W. Kim, and M. H. Ang, "Probabilistic road context inference for autonomous vehicles," in *Robotics and Automation (ICRA), 2015 IEEE International Conference on*. IEEE, To appear.

[14] B. Qin, Z. Chong, S. H. Soh, T. Bandyopadhyay, M. H. Ang Jr., E. Frazzoli, and D. Rus, "A spatial-temporal approach for moving object recognition with 2d lidar," in *International Symposium on Experimental Robotics (ISER)*. Springer, 2014.

[15] S. Ross, J. Pineau, S. Paquet, and B. Chaib-Draa, "Online planning algorithms for pomdps." *J. Artif. Intell. Res.(JAIR)*, vol. 32, pp. 663–704, 2008.

[16] R. Vaughan, "Massively multi-robot simulation in stage," *Swarm Intelligence*, vol. 2, no. 2-4, pp. 189–208, 2008.

[17] Z. Chong, B. Qin, T. Bandyopadhyay, T. Wongpiromsarn, B. Rebsamen, P. Dai, E. Rankin, and M. H. Ang Jr, "Autonomy for mobility on demand," *Intelligent Autonomous Systems 12*, pp. 671–682, 2013.