A Game Theoretic Approach to Decision Making for Multiple Vehicles at Roundabout

Sasinee Pruekprasert[1], Jérémy Dubut[1,2], Xiaoyi Zhang[1], Chao Huang[1] and Masako Kishida[1]

***Abstract*— In this paper, we consider a problem of decision making for multiple autonomous vehicles in a roundabout. In particular, we propose an approach that balances between the safety of all vehicles and their velocity, using a Nash equilibrium to make a decision. Numerical simulations are included to illustrate the effectiveness of the proposed approach.**

I. Introduction

The demand for safety, energy saving, environmental protection, and comfortable transportation services has been increasing. Thus, it is a global consensus to accelerate the development of autonomous vehicles, which incorporate many advanced technologies such as smart sensors and wireless vehicle-to-vehicle communication. For this reason, governments around the world have begun to develop strategies to address the challenges that arise from the autonomous driving [?]. One challenge is ensuring safety and reliability of autonomous driving systems, which is a social and economic problem. The importance of solving this challenge will increase as the degree of automation, and the number of such systems will grow by several magnitudes in the coming future [?], [?].

Roundabouts are often used to improve traffic safety in urban areas. According to different studies, the replacement of signalised intersections by roundabouts reduces injury crashes by 75% [?], [?], and is more suited to where traffic volume is low [?]. From a decision making point of view, they are similar to unsignalised intersections, as they require drivers to decide when to enter, depending on the other vehicles, and influencing the behaviours of the other vehicles.

Although roundabouts are safer than traditional signalised intersections for human drivers, there are still issues to enforce safety for autonomous vehicles. The inner island of the roundabout limits the ability of autonomous vehicles to predict traffic patterns and may lead to a traffic collision. Therefore, critical decision making is the key to collision-free driving in roundabouts. A few game theoretic approaches have shed light on these aspects [?].

In game theoretic approaches, the agents (vehicles) decide their actions by optimising their profit (here, safety and speed) in response to actions of others. This was already observed in the domain of autonomous driving, to model the flow of vehicles in a road with lane change [?], and decision making at unsignalised intersections [?], [?], [?], and roundabouts [?], [?]. In [?], the authors proposed a cooperative strategy in conflict situations between two autonomous vehicles in a roundabout using non-zero-sum games. Each autonomous vehicle tries to shorten its waiting time by analysing possible actions and influences of other vehicles on the game outcome. In [?], the authors proposed to use k-level games for two autonomous vehicles to decide their directions and accelerations.

In this paper, we propose an approach to decision making for several autonomous vehicles in a roundabout. The proposed approach optimises the safety and velocity of vehicles based on Nash equilibria, more specifically by backward induction. This allows a vehicle to decide its control input, using their predictions of other vehicles' behaviours.

The rest of this paper is organised as follows. We first set up the problem in Section II. After describing the flow of our decision making approach in Section III, we focus on the steps that compute the control input and the predictions of the other vehicles' behaviours using game theory in Section IV. In Section V, a set of simulations are carried out to demonstrate the effectiveness of the proposed approach. The section VI concludes the paper.

II. Problem formulation

A. Vehicles in roundabout

This paper considers the decision making of autonomous vehicles at a single-lane w-way unsignalised roundabout intersection. It is assumed that entrances and exits of the roundabout is right-hand traffic and the traffic flows counter-clockwise in the roundabout. A vehicle may enter the roundabout at any entrance and exit at any exit way, but is not allowed to drive backward. The decision making is based on vehicles in the roundabout and those approaching the roundabout, but independent of those already exited the roundabout.

B. Vehicle configurations

In order to focus on the high-level decision making for the autonomous vehicles, this paper does not consider the path trackers and the low-level control layer, which controls the engine to follow a precomputed navigation path (see Fig. 1). For this reason, we consider point-mass vehicle model and assume that the vehicles perfectly follow the given navigation path.

This leaves the only control input for each vehicle i at a time step being its acceleration $a_i(t)$ along the given

The authors are supported by ERATO HASUO Metamathematics for Systems Design Project (No. JPMJER1603), JST.

[1]The authors are with the National Institute of Informatics, Hitotsubashi 2-1-2, Tokyo 101-8430, Japan
{sasinee, dubut, xiaoyi, chao_huang, kishida}@nii.ac.jp
[2]J.D. is also affiliated with the Japanese-French Laboratory for Informatics

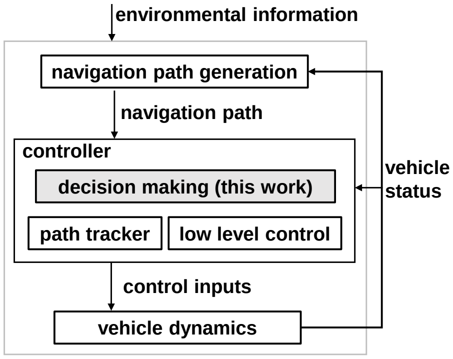


Fig. 1. Motion planning architecture [?], [?] and position of this work.

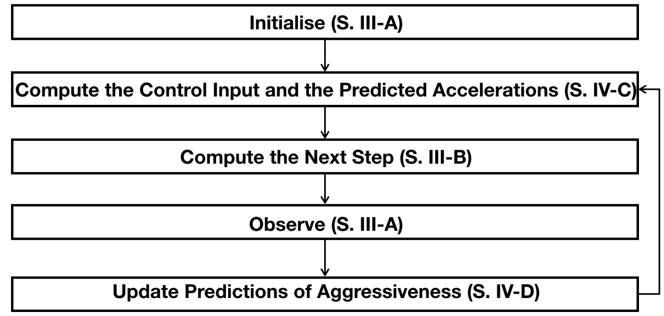


Fig. 2. Outline of the decision making.

navigation path. To simplify the problem, we also assume that each vehicle i chooses an acceleration $a_i(t)$ from a finite set at each time step t, so as to minimise their cost functions.

Let

$$X_i(t) = [r_i(t), \theta_i(t), v_i(t), a_i(t), man_i(t)]^\top \tag{1}$$

represent the configuration of the vehicle i at time step $t \in \mathbb{N}$, where $(r_i(t), \theta_i(t))$ is the position of the vehicle i in polar coordinates, $v_i(t)$ is its velocity, $a_i(t)$ is its acceleration, and $man_i(t) \in \{$"enter" , "inside" , "exit"$\}$ is its manoeuvre, at time step t, respectively. Then, the time-evolution of the configuration of each vehicle i can be represented by

$$X_i(t+1) = F_i(X_i(t), a_i(t)), \tag{2}$$

where the function F_i returns the configuration of the vehicle i after one time step, assuming that $a_i(t)$ is constant between the time steps t and $t+1$.

C. Balancing safety and aggressiveness

The cost functions used to determine $a_i(t)$ balance between two features; safety and velocity. The safety feature is small when the distance with other vehicle is large, while the velocity feature is small when the vehicle's velocity is close to the maximal legal velocity. The balance is determined by the weight vector

$$W_i = [w_i^s, w_i^a], \; w_i^s + w_i^a = 1, \tag{3}$$

where $w_i^s \geq 0$ is the weight of the safety feature and $w_i^a \geq 0$ is the weight of the velocity feature. This weight vector indicates how much the vehicle i values velocity over safety: the higher w_i^a is (and so, the lower w_i^s is), the more aggressive the vehicle will be (and so, the less conservative). Thus, this weight vector determines the *aggressiveness*, which plays an important role in the decision making.

III. Outline of the decision making

This section provides an overview of our approach to decision making summarised below as well as in Figure 2.

First, each vehicle initialises the configuration and the *aggressiveness* for each other vehicle nearby (Section III-A). Then at each time step, each vehicle repeats the followings:

1) Decide the acceleration, by predicting the neighbours' behaviour and minimising its own cost (Sections IV-B and IV-C). This step is the main contribution of the paper.
2) Based on the choice of acceleration and predicted neighbours' behaviour, predict the next configurations (Section III-B).
3) Observe the next configurations (Section III-A).
4) Compare its observations and predictions, and update the predicted *aggressiveness* values of the vehicles nearby (Section IV-D).

A. Observations and initialisation

In the roundabout, each vehicle may not be able to observe all the vehicles, but may observe only the vehicles nearby. Let $\mathcal{N}_i(t) \subseteq \{1, \ldots, m\}$ be the set of vehicles that are being observed by vehicle i at time step t. We assume that $i \in \mathcal{N}_i(t)$, that is, i can observe itself. In our simulations, $\mathcal{N}_i(t)$ is given by the vehicle i itself, the closest two vehicles in front, and the closest vehicle behind, if they are not too far. The observed configuration of the roundabout for the vehicle i at time t is then given by the vector $\widetilde{\boldsymbol{X}}_i(t) = [X_j(t)]^\top_{j \in \mathcal{N}_i(t)}$ of individual configurations (1).

At each time step t, each vehicle i observes its nearby configuration $\widetilde{\boldsymbol{X}}_i(t)$ and computes $\widehat{w}^a_{i,j}(t+1) \in [0,1]$, which is i's prediction for the *aggressiveness* w^a_j of $j \in \mathcal{N}_i(t)$ (to be discussed in Section IV-D). The weight vector $\widehat{W}_{i,j}(t)$ of the vehicle j, predicted by i at time t, is then given by $[1 - \widehat{w}^a_{i,j}(t), \widehat{w}^a_{i,j}]$. For initialisation, let $\widehat{W}_{i,j}(0) = [0.5, 0.5]$ for all $j \in \mathcal{N}_i(0) \setminus \{i\}$ and $\widehat{W}_{i,i}(0) = W_i$.

B. Prediction of the next configuration

Recall that the navigation path generator of each vehicle i provides the function F_i, which computes the next configuration of the vehicle i from its current configuration and a given acceleration a (see Section II-B). Let us consider a given vector of accelerations $\boldsymbol{a} = [a_j(t;s)]_{j \in \mathcal{N}_i(t),\, 0 \leq s < h}$ of vehicles in $\mathcal{N}_i(t)$, for the time horizon from $s = 0$ to $s = h - 1$. Namely, $a_j(t;s)$ is a given acceleration of the vehicle $j \in \mathcal{N}_i(t)$ at the time step $t + s$. We use $\widehat{X}_{i,j}(t, \boldsymbol{a}; s)$ for denoting the configuration of the vehicle $j \in \mathcal{N}_i(t)$ at time $t+s$ predicted by the vehicle i, based on the observation $\widetilde{\boldsymbol{X}}_i(t)$ at time t and the vector of accelerations $\boldsymbol{a}$.

To compute $\widehat{\boldsymbol{X}}_{i,j}(t, \boldsymbol{a}; s)$ at $s > 0$, the vehicle i cannot simply use the function F_j as the vehicle i does not know which path the vehicle j will follow. For example, i does not know which exit j will use. Therefore, the vehicle i predicts a function $\widehat{F}_{i,j}[t]$ using its own function F_i and the observed manoeuvre $man_j(t)$. If $man_j(t)$ is "exit", then i knows that j will use the next exit. If $man_j(t)$ is "enter" or "inside", i computes the navigation path as if the vehicle j stays inside the roundabout indefinitely, possibly turning around several times. We always have $\widehat{F}_{i,i}[t] = F_i$ at each time t as i knows its own navigation path. Then, we define the predicted next configurations $\widehat{\boldsymbol{X}}_{i,j}(t, \boldsymbol{a}; s)$ as follows.

- The predicted configuration at time t is given by the observations:

$$\widehat{\boldsymbol{X}}_{i,j}(t, \boldsymbol{a}; 0) = X_j(t). \tag{4}$$

- The prediction at time $t+s+1$ is computed by assuming that the vehicle j will follow the navigation path given by $\widehat{F}_{i,j}[t]$, from the predicted configuration at time $t+s$, with constant acceleration $a_j(t;s)$:

$$\widehat{\boldsymbol{X}}_{i,j}(t, \boldsymbol{a}; s+1) = \widehat{F}_{i,j}[t](\widehat{\boldsymbol{X}}_{i,j}(t, \boldsymbol{a}; s), a_j(t;s)). \tag{5}$$

IV. Control inputs and predictions, as Nash equilibria

This section is the main contribution of the paper, where we propose a game-theoretic decision making approach to the problem. Since the autonomous vehicles select an acceleration that minimises their costs, the decision making can be naturally viewed as an n-player non-cooperative game played between the vehicles. Using the concept of *aggressiveness*, the decision making can be formulated as a finite perfect-information game.

A. n-player game with perfect information

Let us first briefly review the basic game theory that we use. For details about game theory, see usual textbooks, e.g. [**?**].

We consider a game $G = (P, \Gamma, (H_1, \ldots, H_n))$, where $P = \{p_1, p_2, \ldots, p_n\}$ is the set of *players* (representing the n vehicles), Γ is a finite set of strategies of the players (given by acceleration patterns, Section IV-C) and $H_i : \Gamma^n \to \mathbb{R}^+$ is a cost function (given by safety and velocity, Section IV-B). Each player p_i plays the game by selecting her strategy γ_i from the set Γ, which determines her control input at the current time step. Then, the cost for p_i is given by $H_i(\gamma_1, \ldots, \gamma_n)$. All players aim to minimise their own costs. Let G be a game with *complete information*, i.e., all players know the sets P and Γ, and all functions $H_1, \ldots, H_n$.

A strategy profile $(\gamma_1, \ldots, \gamma_n)$ is called a *best response* of player i if $H_i(\gamma_1, \ldots, \gamma_i, \ldots, \gamma_n) \leq H_i(\gamma_1, \ldots, \gamma_i', \ldots, \gamma_n)$ for all $\gamma_i' \in \Gamma_i$. The strategy profile is also called a *Nash equilibrium* if it is a best response of all players. In other words, a Nash equilibrium is a strategy profile such that no player can reduce her cost by changing her strategy, provided that all other players do not change theirs.

G is a 1-*round sequential game* if there is a total order $\prec$ defined on P, such that the players take turns selecting their strategies according to this order – meaning that if $p_i \prec p_j$ then p_j selects her strategy before p_i – and the game stops after all players have selected their strategies. A sequential game G is a game with *perfect information* if each player remembers the history of all strategies played before her.

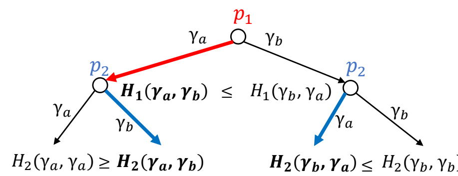


Fig. 3. An extensive-form of a sequential game played between p_1 and p_2, where the set of strategies is $\Gamma = \{\gamma_a, \gamma_b\}$.

The extensive-form of a 1-round sequential game with perfect and complete information can be described as a finite decision tree. Fig. 3 shows an extensive-form of such a game played between two players. As we assume that $p_2 \prec p_1$, p_1 selects her strategy before p_2 at the root of the tree. A Nash equilibrium of the game can be obtain by applying backward induction algorithm on the decision tree. For the game in Fig. 3, a Nash equilibrium can be obtained as following. First, we compute the best responses for p_2 in both subtrees, which represent the cases that p_1 selects γ_a and γ_b. We then compute the best response for p_1 by considering the best responses of p_2 for the subtrees. In this case, (γ_a, γ_b) is a Nash equilibrium.

B. Cost functions

Here, we develop the cost function to be minimised in the decision game. For the whole Section IV-B, we consider $\boldsymbol{X}$ to be a vector of configurations (either observed or predicted), $\mathcal{N}$ indicating a subsets of vehicles (this could be $\mathcal{N}_i(t)$, but not necessarily), $\Phi_i = [\phi_i^{safe}, \phi_i^{velo}]^\top$ to be the feature vector for safety and velocity objectives for the vehicle i, and $W = [w^s, w^a]$ to be a weight vector (typically, the weight vector W_i from Section II-C or given by some predictions $\widehat{W}_{j,i}(t)$ from Section IV-D).

1) Cost at each time step: We first introduce the cost of a vehicle at each time step, which we call the step-cost. The step-cost of the vehicle i is given by

$$S_i(\boldsymbol{X}, \mathcal{N}, W) = W \cdot \Phi_i(\boldsymbol{X}, \mathcal{N}). \tag{6}$$

2) Safety: In order to evaluate the safety, each vehicle i considers the nearest vehicle in front of and the nearest behind within a given distance D along the navigation path (if they exist). Concretely, we consider the pair of vehicles $(f^*, b^*) \in F \times B$ where

$$f^* = \operatorname*{argmin}_f \{\theta_f - \theta_i \mid f \in \mathcal{N} \backslash \{i\}, 0 \leq \theta_f - \theta_i \leq \pi, d(f, i) < D\},$$

$$b^* = \operatorname*{argmin}_b \{\theta_i - \theta_b \mid b \in \mathcal{N} \backslash \{i\}, 0 < \theta_i - \theta_b < \pi, d(i, b) < D\},$$

where θ_j is the angular position of vehicle j and $d(i, j)$ is the distance between the vehicles i and j measured along the navigation path.

Then, we define the safety feature ϕ_i^{safe} by:

$$\phi_i^{safe}(\boldsymbol{X}, \mathcal{N}) = \phi_i^f(\boldsymbol{X}, \mathcal{N}) + \phi_i^b(\boldsymbol{X}, \mathcal{N}), \tag{7}$$

where the features ϕ_i^k, $k \in \{f, b\}$ are defined by:

$$\phi_i^k(\boldsymbol{X}, \mathcal{N}) = \begin{cases} E_{c,en} + \beta_i^k(\boldsymbol{X}, \mathcal{N}) & \text{if } d(k,i) \in [D_{c,en}, D_\infty) \\ & \text{and } man_i = enter, \\ & \text{and } man_k \neq enter, \\ E_c + \beta_i^k(\boldsymbol{X}, \mathcal{N}) & \text{if } d(k,i) \in [D_c, D_\infty) \\ E_\infty + C_\infty.(D - d(k,i))^2 & \text{if } D_\infty \leq d(k,i), \\ \beta_i^k(\boldsymbol{X}, \mathcal{N}) & \text{otherwise.} \end{cases} \tag{8}$$

with

$$\beta_i^k(\boldsymbol{X}, \mathcal{N}) = \begin{cases} 0 & \text{if } k^* \text{ does not exist,} \\ C_{in,k} \cdot (D - d(k,i))^2 & \text{if } man_k = enter, \\ & \text{and } man_i \neq enter, \\ C_{en,k} \cdot (D - d(k,i))^2 & \text{if } man_k \neq enter, \\ & \text{and } man_i = enter, \\ C \cdot (D - d(k,i))^2 & \text{otherwise.} \end{cases} \tag{9}$$

where $C_{in,k} < C < C_{en,k}$, C_∞, $E_c < E_{c,en} \ll E_\infty$, $D_\infty < D_c < D_{c,en}$ are given constants. In (8), ϕ_i^f (resp. ϕ_i^b) is the cost induced by the nearest vehicle in front of i (resp. behind i). All the cases depend on the manoeuvres of f (resp. b) and i. The equations in (9) reflect the fact that, when we are inside the roundabout (the second case), we do not have to care much about the vehicles not entered yet, as we have priority. On the other hand, we have to be extra careful when we are entering the roundabout, as we do not have priority (the third case). The intention of the equation (8) is that a vehicle has to be extra careful when it is close to other vehicles.

3) Velocity: Let v_l be the speed limit of the road. The velocity feature is given by

$$\phi_i^{velo}(\boldsymbol{X}, \mathcal{N}) = \begin{cases} C_{en} \cdot (v_l - v_i)^2 & \text{if } v_l \geq v \text{ and } man_i = enter \\ C_{in} \cdot (v_l - v_i)^2 & \text{if } v_l \geq v \text{ and } man_i \neq enter \\ C_o \cdot (v_l - v_i)^2 & \text{otherwise} \end{cases} \tag{10}$$

where C_{in}, C_{en} (resp C_o) are constant positive coefficients for the cases that v_i is under (resp. over) the speed limit. The intention is that C_o is much bigger than C_{in} and C_{en}, because we cannot allow a vehicle to break the law.

4) Accumulated cost function: We construct the accumulated cost based on the receding horizon control approach [**?**], which determines the control inputs of the vehicles, based on the predicted future up to a horizon time step $h < \infty$. Given a vector $\boldsymbol{a} = [a_j(s)]_{j \in \mathcal{N}, 0 \leq s < h}$ of accelerations (typically, given by a strategy profile, see Section IV-C), the accumulated cost function of the vehicle j, predicted by the vehicle i at time t is:

$$\widehat{K}_{i,j}(t, \boldsymbol{X}, \mathcal{N}, W, \boldsymbol{a}) = \sum_{s=0}^{h-1} \lambda^s \cdot S_j(\bar{\boldsymbol{X}}(t; s), \mathcal{N}, W), \tag{11}$$

where $\lambda \in (0, 1)$ is a fixed discount factor and $\bar{\boldsymbol{X}}(t; s)$ is defined by induction:

- $\bar{\boldsymbol{X}}(t; 0) = \boldsymbol{X}$,
- $\bar{\boldsymbol{X}}(t; s+1) = [\widehat{F}_{i,j}[t](\bar{\boldsymbol{X}}(t; s), a_j(s))]_{j \in \mathcal{N}}$. Recall that $\widehat{F}_{i,j}[t]$ provides the navigation path of j, predicted by i at time t (see Section III-B).

In particular, for $j \in \mathcal{N}_i(t)$, $\widehat{K}_{i,j}(t, \widetilde{\boldsymbol{X}}_i(t), \mathcal{N}_i(t), W, \boldsymbol{a})$ is given by:

$$\sum_{s=0}^{h-1} \lambda^s \cdot S_j([\widehat{\boldsymbol{X}}_{i,j}(t, \boldsymbol{a}; s)]_{j \in \mathcal{N}_i(t)}, \mathcal{N}_i(t), W). \tag{12}$$

Remark that, to predict the accumulated cost of the vehicle j, the vehicle i uses its own observations, in particular, its own $\mathcal{N}_i(t)$. Let us consider the case where j is the nearest vehicle in front of i. There are two situations, depending on whether i can observe its second nearest vehicle in front. If it can, then this vehicle will be in $\mathcal{N}_i(t)$, and then will be considered as the nearest vehicle in front of j in the computation of $\widehat{K}_{i,j}$. If it cannot, this means that this vehicle is far from i, and i will compute $\widehat{K}_{i,j}$, as if there was no vehicle in front of j.

C. Decision making

As introduced in Section IV-A, we employ n-player game to decide the acceleration at each time step. Each vehicle i determines $a_i(t)$ at time t using the Nash equilibrium of the following game $G_i(t)$:

1) The set of players is the set of nearby vehicles $\mathcal{N}_i(t)$.
2) The set of strategies is the set of *acceleration patterns*. More precisely, given a time horizon $h < \infty$, the set Γ of acceleration patterns is then a finite subset of $\mathbb{R}^h$.
3) Given a strategy profile $\mathbf{a} = [a_j(s)]_{j \in \mathcal{N}_i(t), 0 \leq s < h} \in \Gamma^{|\mathcal{N}_i(t)|}$, the cost of the vehicle $j \in \mathcal{N}_i(t)$ is given by:

$$H_j(\mathbf{a}) = \widehat{K}_{i,j}(t, \widetilde{\boldsymbol{X}}_i(t), \mathcal{N}_i(t), \widehat{W}_{i,j}(t), \mathbf{a}) \tag{13}$$

   that is, the accumulated cost of the vehicle j, predicted by the vehicle i. Here $\widehat{W}_{i,j}(t)$ is the weight vector of vehicle j, predicted by vehicle i. (Recall that this vector was initialised in Section III-A, and we will describe in Section IV-D how to update it.)
4) The sequential game is constructed by ordering vehicles in $\mathcal{N}_i(t)$ according to their aggressiveness value, namely, let any total order $\prec$ on $\mathcal{N}_i(t)$ such that if $\widehat{w}_{i,j}^a(t) < \widehat{w}_{i,k}^a(t)$ then $j \prec k$. This means that the more aggressive a vehicle is, the more priority it will have to choose its control input (i.e., its acceleration).

To compute a Nash equilibrium of this game, we use a backward induction from Section IV-A. Namely, a strategy profile is given by a vector $\mathbf{a} = [a_j(s)]_{j \in \mathcal{N}_i(t), 0 \leq s < h} \in \Gamma^{|\mathcal{N}_i(t)|}$ of accelerations, which also includes $a_i(t)$. The predicted configurations for the next h steps are computed as in Section III-B.

D. Updates the prediction of the aggressiveness

1) Updates in a normal situation: In this section, we consider a situation where at least one vehicle from $\mathcal{N}_i(t)$ has a strictly positive velocity.

At each time step t, let $U_i(t) \subseteq \mathcal{N}_i(t) \setminus \{i\}$ be the set of vehicles whose observed configurations $X_j(t+1)$ "differs too much" from the configuration $\widehat{\boldsymbol{X}}_{i,j}(t, \boldsymbol{a}; 1)$ of the vehicle j at time $t+1$, predicted by the vehicle i at time t. For such a vehicle $j \in U_i(t)$, the vehicle i updates the aggressiveness $\widehat{w}^a_{i,j}(t)$ to a value that describes the behaviour of the vehicle j more accurately. Then, vehicle i uses backward induction on several games involving i and j, each game for a w^a ranging in a given set $\mathcal{W}$. More precisely, i considers each game $G_{i,j}(t; w^a)$ defined exactly as $G_i(t)$ in Section IV-C, but replaces $\mathcal{N}_i(t)$ by $\{i, j\}$ (to reduce the computation time) and $\widehat{W}_{i,j}(t)$ by $W = [1 - w^a, w^a]$. Using a backward induction on $G_{i,j}(t; w^a)$ (as in Section IV-C), i computes a strategy profile and predicts the acceleration of j at time $t+1$, which we denote by $\widehat{a}_j(t+1; w^a)$. Let $\widehat{w}^a_{i,j}(t+1)$ be the value w^a for which $\widehat{a}_j(t+1; w^a)$ fits the observed acceleration $a_j(t+1)$ the most closely, that is:

$$\widehat{w}^a_{i,j}(t+1) = \arg\min_{w^a \in \mathcal{W}} |\widehat{a}_j(t+1; w^a) - a_j(t+1)| \quad (14)$$

Vehicle i then update the predicted weight vector by

$$\widehat{W}_{i,j}(t+1) = [1 - \widehat{w}^a_{i,j}(t+1), \widehat{w}^a_{i,j}(t+1)]. \quad (15)$$

Otherwise, if the predicted configuration is close enough to the real configuration (i.e., $j \notin U_i(t)$), then $\widehat{W}_{i,j}(t+1) = \widehat{W}_{i,j}(t)$ and $\widehat{W}_{i,i}(t+1) = W_i$.

2) Updates in a deadlock situation: In this section, we consider the situation where all vehicles in $\mathcal{N}_i(t)$ stop, that is, have zero velocity. This situation is particularly important, as it may induce a deadlock: every vehicle is waiting for the other vehicles to move.

In this case, we want the vehicle i to make a move, as long as it is not in a critical situation. If the vehicle i is not in such a critical situation, we enforce it to make a move by increasing its own aggressiveness. More precisely, we define:

$$\widehat{W}_{i,i}(t+1) = [\max\{0, \widehat{w}^s_{i,i}(t) - 0.5\}, \min\{1, \widehat{w}^a_{i,i}(t) + 0.5\}] \quad (16)$$

and we fix $\widehat{W}_{i,j}(t+1) = \widehat{W}_{i,j}(t)$ for $j \in \mathcal{N}_i(t) \setminus \{i\}$.

V. Experimental results

A. Simulation setup

This section presents simulation results to demonstrate the performance of our decision making. The roundabout and the navigation paths used in the simulation, presented in Fig. 4, are designed based on practical situations [**?**]. All the vehicles drive on the right-hand side of the road and get into the roundabout using one of the four entrances. We consider four types of navigation paths: to turn right (blue dashed line), to go straight (black dashed line), to turn left (red dashed line), and to make a U-turn (green dashed line).

We perform simulations by considering up to five vehicles, whose navigation paths are determined randomly. For the two-vehicle case, the entrance way of each vehicle is

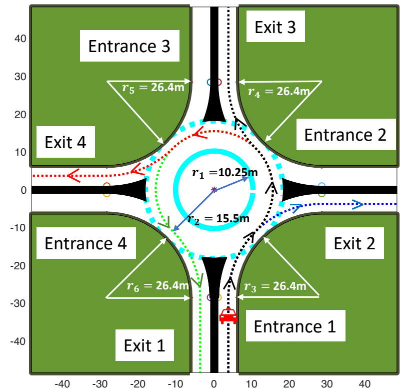


Fig. 4. The navigation paths in the roundabout.

randomised. For the five-vehicle case, one vehicle is placed inside the roundabout at the beginning of the simulation. The initial velocities of the vehicles entering (*resp.* placed inside) the roundabout is randomly selected from $[0, v_l/2]$ (*resp.* $[0, v_l]$), where $v_l = 11\ m/s$ is the speed limit of the road [**?**]. The weight w^a_i of each vehicle i is randomly selected from $\{0.2, \ldots, 0.8\}$.

The decision making is performed every $0.3s$. The observation set $\mathcal{N}_i(t)$ consists of the vehicle i itself, the closest two vehicles in front of i and the closest vehicle behind i, if they are within π radians. For the parameters, we use $\lambda = 0.8$, $D = 15.5\pi\, m$, $C_{in,f} = 2$, $C_{in,b} = 1$, $C = 3$, $C_{en,f} = 6$, $C_{en,b} = 7$, $C_\infty = E_\infty = 10^{200}$, $E_c = 10^{25}$, $E_{c,en} = 10^{35}$, $D_{c,en} = 13$, $D_{c,in} = 10$, $D_\infty = 7$ $C_{en} = 15$, $C_{in} = 0.3$, and $C_o = 10^{15}$.

We use the following patterns of acceleration/deceleration sequences with a time horizon $h = 4$:

- $[-50, -50, -50, -50]$ for a strong break (typically needed for an urgent break),
- $[-20, -20, 0, 0]$ for a small break,
- $[0, 0, 0, 0]$ for no acceleration,
- $[5, 5, 0, 0]$ for a small short acceleration,
- $[5, 5, 5, 5]$ for a constant small acceleration,
- $[20, 0, 0, 0]$ for a strong short acceleration (typically needed when the vehicle has low velocity before entering the roundabout).

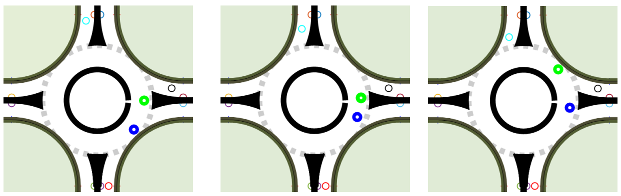
To update the prediction of the aggressiveness (see Section IV-D), a vehicle i updates it's predicted weight $\widehat{w}^a_{i,j}$ if the distance between the observed and the predicted positions of the vehicle j is bigger than $2m$. We use $\mathcal{W} = \{0.1, \ldots, 0.9\}$, although the velocity feature of each vehicle j is initialised within the range $\{0.2, \ldots, 0.8\}$. This is to allow the vehicle i to determine whether $\widehat{w}^a_{i,i} < \widehat{w}^a_{i,j}$ or $\widehat{w}^a_{i,i} > \widehat{w}^a_{i,j}$, so that it can predict the total order $\prec$ (see Section IV-C). A vehicle is considered to be in a critical situation if it is waiting to enter the roundabout, but observes a vehicle that is already inside the roundabout.

We perform 500 simulations for each case. We consider that a collision occurs if we detected two vehicles within $4.5m$. All the programs were coded and run using Matlab 2018a and 2018b.

TABLE I

Simulation Results

No. of Vehicles	Collision Rate(%)	Ave. Minimal Distances	Ave. Total Steps
5	4.2	10.5	107.44
4	2.4	10.4	91.29
3	0	14.2	73.76
2	0	28.5	41.98



(a) (b) (c)

Fig. 5. A dangerous situation at different time steps.

B. Results and analysis

The simulation results are presented in Table I. Columns 1-4 respectively represent the number of vehicles involved in each instance, the collision percentage, the average minimum distances between two vehicles, and the average maximum time step that a vehicles spent in the roundabout.

For the safety feature, we evaluate our simulations by considering the collision rates and the average minimum distance. For two and three vehicles cases, all our simulations (500 each) have ended with no collision and with all vehicles reaching their exits following their navigation paths. For four and five vehicles cases, some collisions were detected (12 and 21 out of 500 simulations respectively).

Another important aspect of our proposed approach is to optimise the velocity and, consequently, the time spent in the roundabout. In average, this is 12.6s (41.98 steps of 0.3s) for the two-vehicle case, while it becomes 32.2s (107.44 steps) for the five-vehicle case.

Fig. 5 illustrates a dangerous situation when two vehicles, blue and green ones, are getting very close to each other. In Fig. 5(a), the green vehicle brakes and waits for the black vehicle to enter the roundabout. In Fig. 5(b), the blue vehicle approaches the green vehicle while expecting the green vehicle to move forward. Both vehicles make rational decisions in this situation. As shown in Fig. 5(c), the blue vehicle brakes and the green vehicle accelerates. Then, the blue vehicle accelerates again after the distance between the two vehicles is large enough. This instance shows that the vehicles can make rational decisions to avoid the potential rear-end collision.

VI. Conclusion

We propose an approach to decision making for multiple autonomous vehicles in a roundabout. Our approach balances between the safety and the velocity feature of the vehicles. Using the proposed concept of vehicles' *aggressiveness*, we formulate the interactions between the vehicles as finite sequential games. The decision making is based on Nash equilibria of these sequential games, which are computed using backward induction. We demonstrate the performance of our approach by performing numerical simulations, showing feasibility and the balance between safety (collision rate, minimum distance) and velocity (average time spent in the roundabout) optimisations. For future works, we will improve our the proposed approach by isolating the scenarios leading to collisions. We will also provide some guarantees that no collisions can occur under some reasonable assumptions of the initial conditions.