AssistMe: Using policy iteration to improve shared control of a non-holonomic vehicle*

Catalin Stefan Teodorescu
Dept. of Electrical and Electronic Engineering
The University of Manchester, Manchester, UK
s.teodorescu@manchester.ac.uk

Tom Carlson
Aspire Create, *University College London*
RNOH, Stanmore, UK
t.carlson@ucl.ac.uk

***Abstract*—An assist-as-needed semi-autonomous control algorithm is designed to address the problem of safely driving a vehicle (a power wheelchair) in an environment with static obstacles. The main idea is to maximize the human driver's control experience while allowing them to navigate safely (the inputs do not lead to collisions). The proposed physically-inspired model-based obstacle avoidance algorithm relies on optimal maps of the expected time for executing a safe stop manoeuvre. These maps are pre-computed using policy iteration in the case of an experienced driver stochastic model. As the burden of complex calculations is handled offline, the online implementation of the algorithm requires little computing resources. Its efficiency is tested experimentally in a study with healthy participants: a statistically significant result confirmed that the proposed algorithm outperforms a baseline rule-based control. A discussion with pros and cons ends this paper.**

***Index Terms*—shared control, policy iteration, obstacle avoidance, medical robotics**

I. INTRODUCTION

This work is inspired by a real need for bespoke control solutions applicable to medical robotics. In the European project "ADAPT", clinicians (medical doctors) identified the needs of wheelchair users [1] and in response, roboticists proposed to make use of technology and novel scientific approaches to tackle those challenges, one of which is addressed in this article.

A. The need for assistive technology

In health and social care, about 840,000 people in the UK are regular wheelchair users [2] living with multiple impairments (physical, cognitive, visual, hearing) who might benefit from an assisted robotic wheelchair that can help them navigate safely in any indoor or outdoor environment. There are currently no such commercially available devices on the market. Consequently, they are at risk of bumping unintentionally into obstacles, hurting themselves or the people around them.

Over the last half century, the UK has seen an increase in life expectancy and this trend is expected to continue [3]. To meet the needs of an ageing society, research in innovative assistive robotic technologies has the potential to empower healthy aged people towards living an independent and active life, while keeping them safe and socially connected [4], [5], thus increasing their self-esteem. The beneficiaries of such technology account for 58% of the people aged 60 and over in UK who suffer from mobility difficulties [6]. In 2014 there were 14.9 million people aged 60+ and their number is expected to grow to 18.5 million by 2025 [7].

B. Shared control

Assist-as-needed [8] or shared control [9] is a concept involving collaboration between a human and a machine. The human operator informs the machine about their objective (e.g. by actuating a joystick) and the machine interprets that action, then implements it in an optimal way. The algorithms studied in the literature can be divided in two categories, depending on the need to model system dynamics. On the one hand there are the *black-box* control solutions like the Probabilistic Shared Control [10], optimal control in [11], [12], and Dynamic Window Approach [13], task-oriented control [5]. Arguably the most interesting feature of these algorithms resides in their generic value: they are most suitable in situations where models of the wheelchair dynamics or the user intention do not exist, and can easily be switched (adapted) from one vehicle type to another one. The second category of algorithms rely on *white-box* control solutions that require model identification; discussed next.

C. Model-based control

Moving vehicles obey physical laws which are well-known and understood (e.g. inertia, gravity, etc.). Consequently, it is natural to create a model integrating the capabilities of the machine. Although models in general are a mere reflection of reality, control-oriented models permit taking into account the uncertainties from the early phase of designing the control. The challenge is to be able to derive control policies based on the control-oriented model and the limited unpredictable information coming from: (i) the sensors observing the environment (e.g. low-cost time-of-flight sensors that can measure only the 1D distance to obstacles whereas their 3D shape remains unknown), and (ii) the human user (e.g. the joystick interface is a projection of what the user thinks, and their intention). In what follows, we shall make the distinction between the availability of significant amount of sensor data (e.g. measurements at high rate) and the limited information provided by each sensor (e.g. 1D distance to an obstacle is not sufficient to know the precise location or form of an obstacle in 2D Cartesian space; 2D joystick data is only informative about the immediate intention of the driver, not the long-term intention like getting closer to a table in

*This work was funded by the project INTERREG VA FMA ADAPT – "Assistive Devices for empowering disAbled People through robotic Technologies" http://adapt-project.com/index.php. The FMA Program is a European Territorial Cooperation program which aims to finance ambitious cooperation projects in the border region between France and England. The Program is supported by the European Regional Development Fund (FEDER).

order to serve their lunch). To compensate for this limited information available via online measurements, recent state-of-the-art research uses stochastic models.

D. A user-centered design

Taking into account the intention of the user is essential. An assistive machine should complement their intention in such a way that the user feels empowered and can make use of their existing functionality (e.g. physical, cognitive, visual, hearing, etc.) [14]. Contrary to able-bodied people who might be more in favour of full automation (think of Tesla cars driving autonomously), the potential users of semi-autonomous assist-as-needed technologies want to maximise their level of control, and be able to develop new skills. They favour technology that empowers the user rather than further disabling them by taking away control authority and treating them like a precious piece of cargo. Therefore, instead of relying on generic control algorithms intended to satisfy all users, we shall put forward a method that can be tailored to specific groups of users, and in this paper we consider the example of experienced drivers. This is done by taking into account the stochastic model of the particular user group (experienced drivers) from the early phase of designing the assist-as-needed control. The same methodology could be followed for different user groups, e.g. novice, or those with particular pathologies.

E. Optimal control

One of the challenges faced by implementing Probabilistic Shared Control algorithms on smart wheelchairs is the limited computational resources available for online processing at a minimal cost [10]. Conversely, Stochastic Dynamic Programming (SDP) is a control design technique that can handle a stochastic model by solving a global optimization problem under constraints. The solution is not analytical, instead it is computed numerically, often on high-performance computers able to handle the burden of intensive computations and fairly large matrices, and comes in the form of optimal control policies [15]. These operations are carried out offline using algorithms like *policy iteration* and the outcome, namely a lookup table containing the optimal actions, is implemented online on hardware with limited computational resources (e.g. industrial controllers or single-board computers). This approach is particularly appealing to the automotive sector [16] and we propose to translate it to smart wheelchair applications.

F. The control problem

In this article we address the safe-driving problem of a *semi-autonomous* vehicle (the power wheelchair in Fig. 1) in an environment with static obstacles, as depicted in Fig 2. In Fig. 2, the angle $\alpha_b \geq 0$ defines a sector area $[-\alpha_b, \alpha_b]$ or field of view where ultrasonic sensors detect obstacles. The goal is to enable safe navigation while roaming freely between obstacles. There is no fixed end destination location. Note the *semi-autonomous* vehicle case is still an active area of research [12], unlike the purely *autonomous* vehicle case, which is out-of-scope for the purpose of this article and where robust and well-established path planning algorithms are readily available [17], [18, §11.1.5].

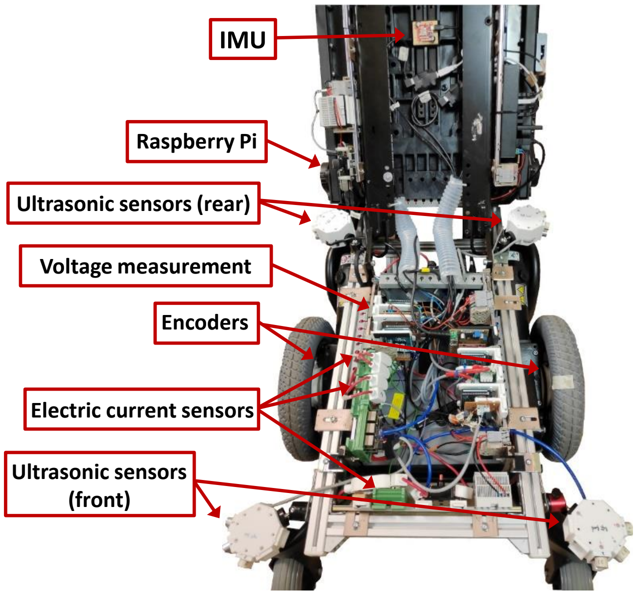


Fig. 1: View under the seat of an off-the-shelf commercial power wheelchair converted into a prototype instrumented semi-autonomous robotic vehicle, where bespoke electronics was added

G. Our contribution

In this article we propose a novel obstacle avoidance shared control algorithm based on policy iteration. To the best of our knowledge, policy iteration has not been proposed yet in the scientific literature for addressing the problem stated in section I-F. It is our intention to make a significant contribution in this setting as follows.

Inspired from the automotive sector, where a rather limited number of successful implementations of policy iteration algorithms have been published (see, e.g. [19], [20]), this article builds upon the seminal work [21]. Our article is similar in the sense that we shall reuse the same idea of implementing policy iteration to compute global optimal policies in conjunction with a stochastic model. However, our contribution differs with respect to the following points: (i) the application is different (a power wheelchair, not a passenger car); (ii) the stochastic model is different (we use the experienced driver model, not a driving cycle model); (iii) the level of desired autonomy is different (control of semi-autonomous vehicles, not autonomous vehicles); (iv) the optimization is carried out in terms of total time to complete the task (obstacle avoidance), not energy (minimizing fuel consumption *versus* battery usage in hybrid electric vehicles along given driving cycle models).

II. Policy Iteration

In this section we formulate an infinite horizon time-optimal stochastic shortest path problem [15, §7.2]. This is a finite-state problem involving a finite (but rather large) state space. The framework was largely inspired by the inventory control problem in [15, Ex. 1.3.2, p. 28]. The modeling of the vehicle dynamics is presented first, followed by the driver model and lastly the results of implementing policy iteration are shown. The central part of this section is the computation of a map showing the average time it takes our

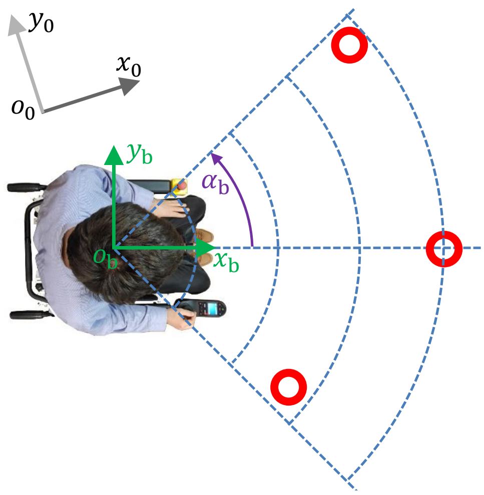


Fig. 2: View from above: instrumented wheelchair advancing forward in a two-dimension Cartesian environment with obstacles (in red). $o_0x_0y_0$ is the fixed (inertial) frame and $o_\mathrm{b}x_\mathrm{b}y_\mathrm{b}$ is the moving (base) frame

TABLE I: Nomenclature

Symbol	Meaning	Units
α_b	angle of observation	rad
x_k	position of the vehicle expressed in coordinates of the fixed frame 0 in Fig. 2	m
d_max	maximum distance an obstacle can be detected (perceived) by a ultrasonic sensor, expressed in coordinates of the moving base frame in Fig. 2	m
v_k	linear velocity of the vehicle	m/sec
v_min, v_max	minimum and maximum linear velocity, respectively	m/sec
$v_{\mathrm{d},k}$	driver's demand (or intention) expressed by actuating the joystick	m/sec
$v_{\mathrm{u},k}$	algorithmic control variable acting additive to the joystick signal	m/sec
Δt	sampling time	sec
σ_1, σ_2	physically-inspired experimentally identified parameters [24]	–

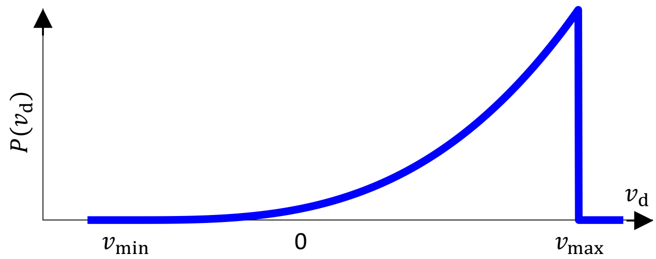


Fig. 3: Conceptual illustration of the probability distribution of the experienced driver model

experienced driver model to execute a safe stop manoeuvre. This information is used later on in section III to elaborate the assist-as-needed semi-autonomous control algorithm.

A. Vehicle model

Various wheelchair dynamical models are proposed in the literature, e.g. [22]. Building on our previous work [23], the vehicle in Fig. 1 advancing in straight line is modeled as a point mass governed by the following dynamics:

$$x_{k+1} = \max(0, \min(d_\mathrm{max},\, x_k + v_k \Delta t)) \tag{1a}$$

$$v_{k+1} = \max(v_\mathrm{min},\, \min(v_\mathrm{max}, \sigma_1 v_k + \sigma_2(v_{\mathrm{d},k} + v_{\mathrm{u},k})) \tag{1b}$$

with the constraint that $v_{\mathrm{u},k}$ is such that

$$v_\mathrm{min} <= v_{\mathrm{d},k} + v_{\mathrm{u},k} <= v_\mathrm{max} \tag{2}$$

All variables and parameters are defined in Table I together with their units of measure; subscript k indicates the time index. The saturation in (1) ensures the existence of a finite number of states and is necessary for implementing the policy iteration algorithm. Furthermore, it has physical meaning:

- in (1a), the space-horizon of the vehicle motion starting at $x_k = 0$ is limited to the maximum viewing range d_max of the ultrasonic sensors in Fig. 1;
- in (1b), a low-level safety functionality on each servo-drive actuating the motors ensures the linear velocity is kept above $v_\mathrm{min} < 0$ and below $v_\mathrm{max} > 0$. This bypasses (overrides) any requested control action $v_{\mathrm{d},k} + v_{\mathrm{u},k}$.

Next, the physical meaning of the constraint (2) is to avoid requesting an input velocity $v_{\mathrm{d},k} + v_{\mathrm{u},k}$ that is more than the wheelchair's power module profile is set to handle.

B. A stochastic driver model

We carried out a preliminary analysis of driver behaviour which showed that experienced drivers have a tendency to predominantly use maximum velocity when advancing forward. This observation motivated the model shown in Fig. 3. For the purpose of the policy iteration algorithm implemented in this article, and without loss of generality, in what follows we shall use the generic stochastic user model depicted in Fig. 3. Future experimental studies could fine tune it and generate different models for different types of users, capturing their preferred driving styles.

C. Implementation and Results

Policy iteration is implemented in this section, on the simplified case $\alpha_\mathrm{b} = 0$ in Fig. 2, of a vehicle advancing in straight line with an obstacle located in front of it. The results of this section are used in section III where a solution to the control problem from section I-F is presented, covering both the situation $\alpha_\mathrm{b} = 0$ as well as $\alpha_\mathrm{b} > 0$ in Fig. 2, of a vehicle able to detect multiple forward-facing obstacles on the sector area $[-\alpha_\mathrm{b}, \alpha_\mathrm{b}]$ in Fig. 2.

Using the stochastic driver model from section II-B, we are now interested in determining the minimum expected *time to termination*, defined as the time elapsed from the initial state (x, v) of the moving vehicle to the moment where it stops safely as close to that obstacle, but without hitting it. We call this a safe stop manoeuvre. Formally, we address an optimization problem where the aim is to compute the total cost:

$$J(x, v) = \min_{v_{\mathrm{u},k}} \mathop{\mathbb{E}}_{v_{\mathrm{d},k}} \left\{ \sum_k g(x_k, v_k) \right\} \tag{3}$$

along trajectories defined by (1) and such that constraint (2) holds; $k = 1, .., \infty$ is the number of stages. In (3), $\mathbb{E}$ is the *expected value* operator expressing a sum of all possible (feasible) trajectories that are weighted by their probability of occurrence. The policy iteration algorithm that implements the cost (3) will compute numerically that optimal control policy which works best for all feasible trajectories (in the sense of a weighted sum). Intuitively, the dominant trajectories that have higher weights will account the most in the process of computing the optimal control policy. The cost per stage $g(\cdot,\cdot)$ in (3), is defined next.

1) Cost per stage without penalty: First, we defined an environment similar to a toy circuit where a toy vehicle bumps into rubber damper walls physically limiting its ability to advance any further (beyond the boundaries of the circuit) and also protecting the toy from damage. These boundaries for states (x_k, v_k) are defined by the saturation limits in (1), and placing them on a finite grid is necessary for computational tractability reasons. Choosing

$$g(\cdot,\cdot) \equiv \Delta t \tag{4}$$

in (3), the control policy comes in the form of a bang-bang control, typical solution when solving time-optimal problems [25]. It represents an aggressive maneuver of that specific stochastic driver model advancing in minimum time towards the obstacle, but still being able to stop safely near the obstacle, without bumping into it. An interesting fact of this formulation is that the total cost J in (3) has a physical meaning: the map in Fig. 4a shows the average elapsed time, measured in seconds, for executing this safe-stop maneuver by the stochastic driver model. Intuitively, the slope of the map in Fig. 4a descends as the vehicle starts closer to the terminal state.

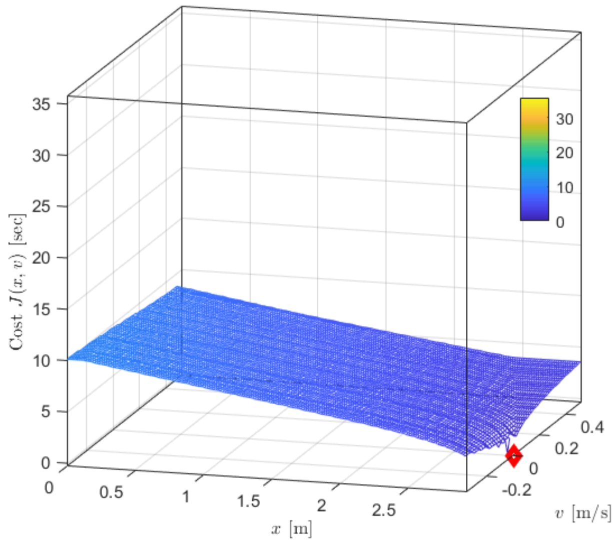
Second, in the next section we show how to remove the artificial limitation imposed by using this toy circuit concept. For that, note that bumping into the rubber damper walls can only occur in two situations: (i) $x = 0$ and $v < 0$; (ii) $x = d_{\text{max}}$ and $v > 0$. We introduce soft constraints [25] in those two situations.

2) Cost per stage with penalty: By choosing

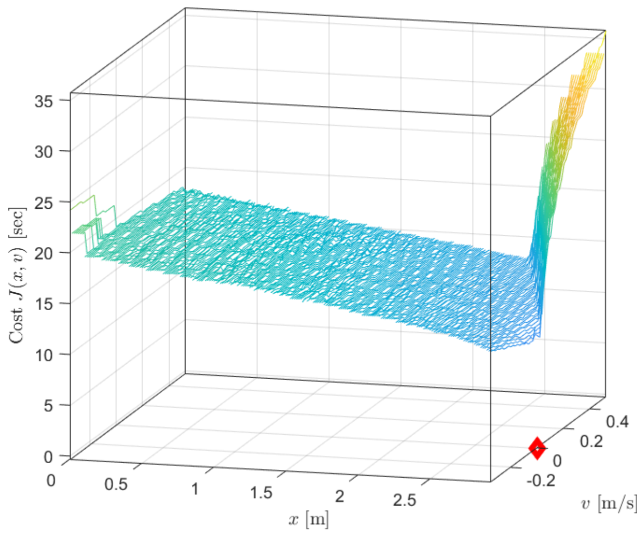
$$g(x_k, v_k) = \begin{cases} \beta\Delta t, & \text{if } (x_k = 0 \text{ and } v_k < 0) \text{ OR} \\ & \quad (x_k = d_{\text{max}} \text{ and } v_k > 0) \\ \Delta t, & \text{otherwise} \end{cases} \tag{5}$$

with $\beta \gg 0$ a penalty factor, the cost function in (3) changes both its mean value as well as the shape at extremities: see Fig. 4b. First, we analyze the shape. Interestingly, we can now easily identify those initial states (x, v) starting from which the vehicle would crash into the obstacle (the area around $x = d_{\text{max}}$ and $v > 0$), or would advance backwards beyond the range of sight of the sensors ($x = 0$ and $v < 0$): they correspond to regions where the cost function surges in Fig. 4b. Those regions act as restrictions where assistive control should not be enabled, instead safety measures need to be enforced (e.g. by taking authority from the user, active braking, etc.) which is a topic out-of-scope for this article.

Second, we analyze the values of the map in Fig. 4b. A drawback of having introduced the penalty in (5) is that the mean values of the maps in Fig. 4b went up by approximately 150%, in average, compared to the results in Fig. 4a, making them lose their original physical meaning (namely, the average time for executing a safe stop maneuver). The results of this section are carried on to the next one, where we explain the assistive control algorithm.



(a) without additional penalty on the cost per stage $g(\cdot)$ from (4)



(b) with additional penalty on the cost per stage $g(\cdot)$ from (5)

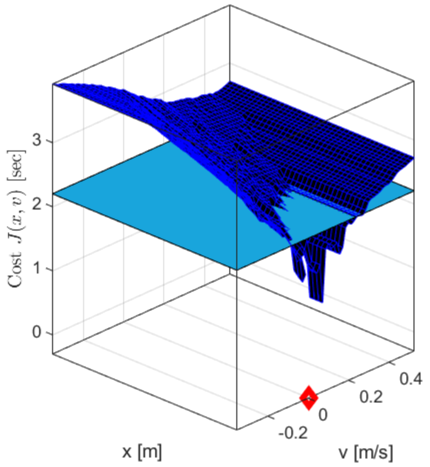
Fig. 4: Policy iteration results. $J(x, v)$ represents the minimum expected time to reach the terminal state (in red), for the experienced driver stochastic model

III. Assist-as-needed semi-autonomous control (ASC)

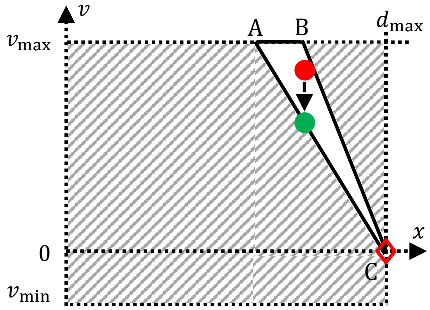
The proposed ASC algorithm is designed in two steps. First, the linear velocity control follows from the map computed in the previous section II. Second, the angular velocity control is adjusted proportionally to the reduction in the linear velocity.

A. Linear velocity control

We start by analyzing the simplified case $\alpha_{\text{b}} = 0$ in Fig. 2. An arbitrary threshold representing the average time to execute a safe stop maneuver for our stochastic driver model is set to, say, 2.2 sec as illustrated by the horizontal plane in Fig. 5a. The other element of Fig. 5a is the map taken from Fig. 4a. The projection onto the (x, v) state space of the intersection between this plane and the map, gives a boundary, conceptually illustrated by the line segment



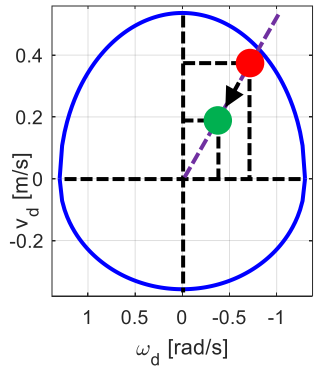
(a) Setting up an arbitrary threshold value of 2.2 seconds for the time to execute a safe stop maneuver (in cyan), in relation to the computed map (in blue) from Fig. 4a (zoom in around the terminal state in red)



(b) Conceptual drawing in state-space of the working principle of the ASC algorithm for controlling the linear velocity: the idea is to reduce any unsafe linear velocity (see e.g. the red solid circle) within the area of the triangle ABC towards a safe value (see the green solid circle); note the terminal state is illustrated as a red rhombus

Fig. 5: Designing the ASC algorithm

(a) joystick device



(b) corresponding joystick plane

Fig. 6: Working principle of the ASC algorithm for controlling both the linear velocity and the angular velocity: the unsafe user demand (say, the red solid circle on both images) is reduced linearly (proportional) to a safe value (say, the green solid circle)

AC in Fig. 5b. In addition to that, the line segment BC in Fig. 5b stands for the boundary where map values surge in Fig. 4b. Note the line segment AB expresses a physical constraint where the vehicle advances at maximum velocity $v_{\max}$. Finally, the triangle ABC in Fig. 5b defines a region where it is meaningful to enable ASC as a way to protect the driver from crashing into the obstacle.

Motivated by the observation that $v = v_k = v_{\rm d} + v_{\rm u}$ in steady-state in (1a), the guiding principle for driving the vehicle assisted, is to use $v_{\rm u}$ as means to decrease v_k in situations where the user demand $v_{\rm d}$ is too high in spite of the obstacle being too close. Specifically, measured values v inside the triangle ABC in Fig. 5b (e.g., the red solid circle) need to be lowered, to reach the safe boundary depicted by the line segment AC in Fig. 5b (e.g., the green solid circle).

Next, we show how to control the angular velocity in the more general case $\alpha_{\rm b} > 0$ in Fig. 2 of a vehicle advancing in two-dimension Cartesian space, with multiple forward-facing obstacles.

B. Angular velocity control

In case of multiple obstacles detected ahead (e.g. three obstacles in Fig. 2), the ASC handles the most imminent danger (the closest obstacle).

To complete the proposed ASC algorithm we need to also consider the angular velocity ω_k. In steady-state, $\omega_k = \omega_{\rm d} + \omega_{\rm u}$, where $\omega_{\rm d}$ represents the input from the user (see the horizontal axis in Fig. 6) and $\omega_{\rm u}$ is the contribution of the assistive control. There is a specific correspondence between the position (location) of the joystick in Fig. 6a and one point on the joystick plane in Fig. 6b: see, e.g. the red solid circle on both images in Fig. 6. Physically, that point in Fig. 6b expresses the velocities reached in steady-state by the vehicle, in the absence of assistive control ($v_{\rm u} \equiv \omega_{\rm u} \equiv 0$). In other words $v_k \to v_{\rm u}$ and $\omega_k \to \omega_{\rm u}$ as k progresses. For more details, see e.g. [23].

The idea put forward in angular velocity control is to lower ω_k proportionally to the reduction in the linear velocity control according to section III-A. To visualize this, the red solid circle in Fig. 6 corresponds to an unsafe user demand $(v_{\rm d}, \omega_{\rm d})$ and consequently it is reduced towards a safe value illustrated by the green solid circle, which corresponds in steady-state to velocities $(v_k = v_{\rm d} + v_{\rm u}, \omega_k = \omega_{\rm d} + \omega_{\rm u})$.

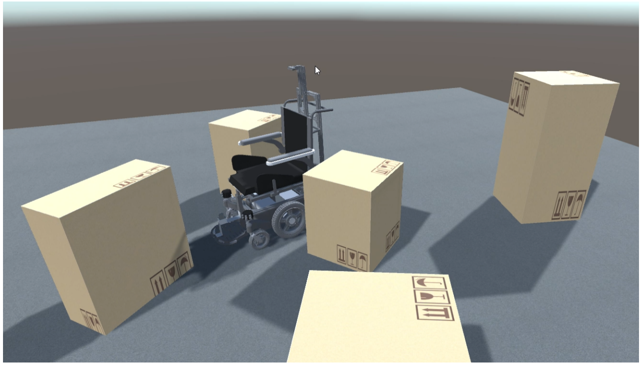
IV. Experiments

To validate the proposed ASC, we used the research platform in Fig. 1. To build the environment awareness capability for the robotic vehicle we used: (i) the standard joystick to capture the user's input $(v_{\rm d}, \omega_{\rm d})$, (ii) the wheel encoders to estimate the actual velocity v_k, and (iii) the front-facing ultrasonic sensors to detect the nearby obstacles (see Fig. 1). The software architecture relies on a single-board computer (a Raspberry Pi 3b+) running Robot Operating System (ROS) on Ubuntu 16.04. The C++ code used for our experiments is available on https://github.com/UCL-Aspire-Create/whc_pub_ASC_policy_iter.git; Fig. 7a shows the lab environment where experiments took place, and the simulators in Figs. 7b and 7c were used for visualization.

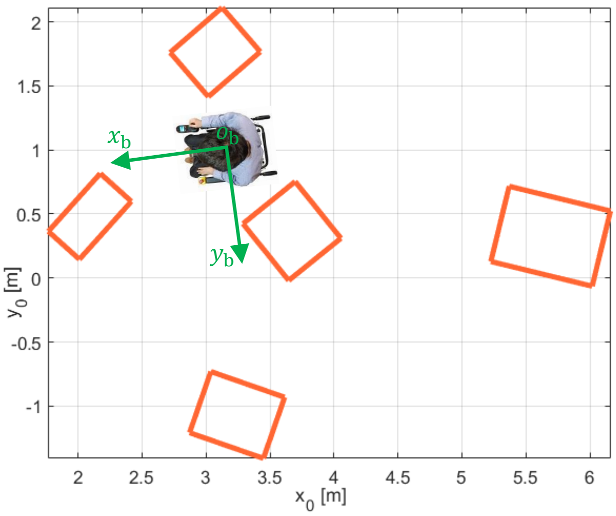
A. A study with participants

Due to Covid restrictions, we were only permitted to recruit four participants, all able-bodied healthy volunteers,

(a) real world



(b) real-time 3D simulation in Unity using Optitrack markers placed on obstacles and the wheelchair



(c) offline 2D simulation in Matlab (view from above); note the obstacles illustrated in orange

Fig. 7: The circuit used for testing the obstacle-avoidance algorithms: in real world and simulated

who were experienced in driving the wheelchair. Each participant was initially shown the circuit in Fig. 7a with 5 cardboard box obstacles, then asked to follow this protocol. They needed to complete 4 back-to-back iterations (repetitions), each consisting of the same succession of operations: (1) start driving the wheelchair from an initial position located to the left-hand side of Fig. 7a; (2) drive between the obstacles; (3) pick one sticker attached to the cardboard box located to the right-hand side of Fig. 7a; (4) drive back between the obstacles to the starting point; (5) place the sticker in a basket. At the end of the fourth iteration, 4 blue stickers are counted inside the basket. Note that drivers were only allowed to advance forward (no reversing). A video recording is available on https://youtu.be/X-dQoLpYquQ showing one participant driver completing one iteration of the circuit. This study is organized as a competition between the participants: the winner is the person who can drive in minimum time and with the least number of collisions. Say t_d is the time to complete the track and c_d is the number of collision, then each participant receives a score s representing a *penalized task completion time* measured in seconds, based on their performance:

$$s = t_\mathrm{d} + \alpha\, c_\mathrm{d} \tag{6}$$

where $\alpha > 0$ acts as a penalty for hitting obstacles. Inspired from circuit racing [26], we use $\alpha = 5$ sec.

The aim of this study was to assess whether the ASC brings any benefit compared to: (i) the situation with no assistance (NA): this is used as the *control group* [27]; (ii) a baseline assist-as-needed semi-autonomous control: here we shall use our previous rule-based (RB) algorithm [28]. Hence, we address the *research question*: how does the driving condition (NA, RB, ASC) influence the user performance? We make a first hypothesis that there is a difference between all these three driving conditions (NA, RB and ASC), and second that ASC is superior to RB. In order to reduce order effects (structural bias), the order of the driving conditions (NA, RB, ASC) is assigned randomly for each participant using *simple random sampling* [29, §3]. An ethics approval was granted for this study by UCL Research Ethics Committee (ref. 6860/011).

B. Data collection and visualization

The table in Fig. 8a shows the collected data: each row represents a single participant's data; each column corresponds to one of the three driving conditions under consideration (NA, RB and ASC). This data converted into scores using (6) can be visualized in Fig. 8b. The mean values (red lines in Fig. 8b) suggest that the NA group performs the best, followed by the ASC group and lastly by the RB group. Before drawing any conclusion, we need to check the statistical significance of this observation using inference statistics.

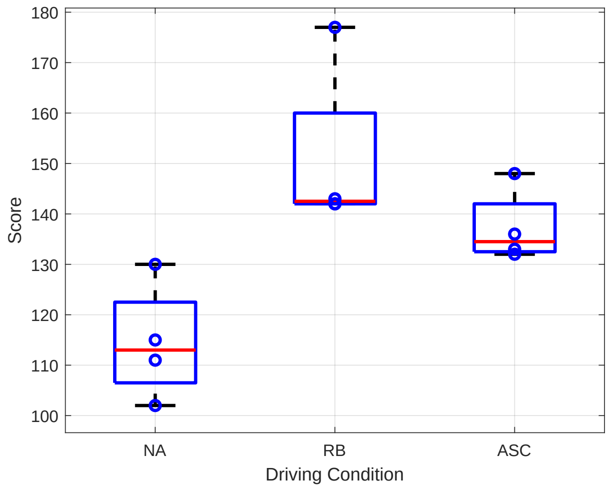
C. Statistical analysis

In order to compare scores of the 3 aforementioned groups, we shall use ANOVA [27], [29], [30]. Before deciding which ANOVA variant to use, first we checked the assumption of parametric data. Only one group, NA, meets this assumption (Shapiro-Wilk $p = 0.836$, skewness $\rho = 0.724$ and no outliers), whereas the other two groups violate this assumption, RB (Shapiro-Wilk $p = 0.003$, skewness $\rho = 2$ and no outliers) and ASC (Shapiro-Wilk $p = 0.125$, skewness $\rho = 1.7$ and no outliers). This motivates the use of the One Way Repeated Measures Friedman's ANOVA which is a non-parametric alternative to the more popular parametric F-test.

Hence, as the data violated the parametric assumption, a Friedman's ANOVA – a *within participants* design [30],

Participant	NA	RB	ASC
1	111 (0)	143 (0)	132 (0)
2	125 (1)	142 (0)	136 (0)
3	102 (0)	142 (0)	133 (0)
4	110 (1)	172 (1)	148 (0)

(a) Data collection. Time to complete the circuit in seconds, and number of obstacles hit (indicated between round brackets)



(b) Visualizing the data. Box plots of the scores (lower is better)

Fig. 8: Experimental *sample data* for the 3 different groups: NA = no assistance; RB = rule-based control; ASC = Assist-as-needed control

was conducted to test whether the driving condition (NA, RB or ASC) affects the participants score in completing the circuit track. The analysis revealed a significant main effect of driving condition ($\chi^2(2) = 8$, $p = 0.018$). The post-hoc analysis using pairwise comparisons (Durbin-Conover) revealed a significant difference ($p < .001$) between pairs of groups: (i) the NA (mean=115, $\sigma = 11.7$) and RB (mean=151, $\sigma = 17.3$); (ii) the NA and ASC (mean=137, $\sigma = 7.37$); (iii) the RB and ASC. This suggests that ASC is superior to RB although both are outperformed by NA.

D. *Limitations of the study*

Note that increasing or decreasing the arbitrary value for the penalty α on hitting obstacles, will have an immediate effect on the scores and thus the entire statistical analysis. For instance, inspired by and consistent with circuit racing practice [26], another option would be $\alpha = 10$ sec. Alternatively, more general and elaborate methods to adjust the weighting factor α are discussed in [31].

Due to the ongoing coronavirus pandemic, we were only able to recruit a rather limited number of participants for this study. Although results are statistically significant, these should be interpreted with caution and a follow-up study involving more participants is recommended.

E. *Discussion*

The rather intriguing fact that by disabling assistive control (NA group) the participants in this study obtained better scores compared to enabling assistive control (RB and ASC groups) can be explained by multiple factors. First, by the nature of these assistive control solutions: both gradually take authority from the driver as the risk of collision increases. This translates into reduced effective velocities and thus the time to complete the circuit increases. Second, the participants in this study were experienced drivers (although not experts), capable of handling the obstacle avoidance task rather easily (NA group). Third, participants were all able-bodied, whereas we expect assistive control to prove its real-world usefulness for people living with impairments (e.g. cognitive, visual, hearing, physical, etc.) [1]. To summarize, there is a trade-off between the risk of collision and the ability to finish the circuit quickly: the price to pay for reducing the risk of collision when enabling assistive control is to tolerate (accept) that it takes more time on average to complete the circuit. From a broad perspective, our contribution has the potential to transform the lives of many people by creating a transportation system that empowers the user by means of technology to carry out everyday tasks safely, without bumping unintentionally into static obstacles.

Our statistical analysis used a combined metric (a score), by weighting together the task completion time and the number of collisions. If instead, we are interested in analyzing each metric separately. The following results were obtained. First, the analysis using the task completion time revealed a similar (analogous) result: a significant effect of driving condition ($\chi^2(2) = 8$, $p = 0.018$). Second, the analysis using the number of collisions showed a non-significant effect of diving condition ($\chi^2(2) = 3$, $p = 0.223$), meaning that we cannot conclude that ASC is superior to NA or RB: the fact that no collision occurred for ASC group in Table 8a, contrary to the other groups, might be the result of a lucky sample; we need more participants to the study in order to obtain a statistically significant result in this setting.

Instead of using *policy iteration*, a similar result could have been achieved using the less computationally-intensive *value iteration* algorithm [15, §7.2]. However, we chose to implement *policy iteration* in view of future research, where we intend to make use of the computed optimal control actions.

An interesting fact of this semi-active (dampening) control design is that in all situations it will not contradict the user's intention (e.g. cases where the driver intends to advance to the left but the SC would make the vehicle turn right, would not occur). This is due to the linear (proportional) relation between the locations of the unsafe demand and the corrected one, as we have explained in section III, say the red solid circle and the green solid circle in Fig. 6. Such conflicting situations were analysed theoretically in [32] and this paper provides a practical solution to that.

The statistical analysis suggested that ASC outperforms RB, which is quite a positive result because it shows a progress in our research. However, the scope of this result is limited to a particular circuit and does not account for other real-world scenarios like the standardized circuits in [1] (e.g. entering an elevator, advancing in a wide corridor with static obstacles, etc.). Follow-up studies are necessary to address these situations. Another challenging situation is avoiding dynamic obstacles like pedestrians walking [33]. This would require an extension of the ASC algorithm, and a possible way to do this is to start from the observation that the situation of a moving vehicle plus a moving obstacle is equivalent to the situation of a (faster) moving vehicle and a static obstacle.

V. Conclusions

In this article we developed and tested an assist-as-needed algorithm (we called it AssistMe) that provides a robotic vehicle with the intelligence (or capability) to avoid obstacles by collaborating with the user and thus ensuring a safe driving experience. We showed how this algorithm can be personalized to meet the needs of a specific stochastic user model (e.g. the experienced driver model). The algorithm makes use of pre-computed optimal maps of the average time to hit an obstacle located ahead of the vehicle: we formulated a time-optimal stochastic shortest path problem and solved it numerically by implementing a computationally tractable policy iteration algorithm. To test it experimentally, low-cost hardware components are required (ultrasonic sensors, wheel encoders, a single-board computer) mounted on a powered wheelchair. An experimental study with healthy participants, showed a statistically significantly improved score of this assist-as-needed algorithm over a baseline rule-based control.

Acknowledgment

Authors would like to thank Marie Babel and her team (in particular, François Pasteau and Louise Devigne) at INSA and IRISA in Rennes, France, for providing us with an RNET controller compatible with the Salsa M2 wheelchair in Fig. 1, used to: (i) publish standard joystick data on a ROS topic, and (ii) send *virtual joystick* messages on the RNET bus to the power module. We acknowledge the work of George Walker for developing the Unity environment in Fig. 7b.

References

[1] E. Leblong, B. Fraudet, L. Devigne, M. Babel, F. Pasteau, B. Nicolas, and P. Gallien, "SWADAPT1: Assessment of an electric wheelchair-driving robotic module in standarized circuits: A prospective, controlled, and randomized pilot study," *J. of NeuroEngineering and Rehabilitation*, vol. 18, no. 140, pp. 1–12, 2021.

[2] National Health Service (NHS), "Improving wheelchair services," accessed on April 4, 2022. [Online]. Available: https://www.england.nhs.uk/wheelchair-services/

[3] C. Thompson, "Meeting the challenges of an ageing population," National Health Service (NHS), accessed on April 4, 2022. [Online]. Available: https://www.england.nhs.uk/blog/catherine-thompson/

[4] L. Devigne, M. Aggravi, M. Bivaud, N. Balix, C. S. Teodorescu, T. Carlson, T. Spreters, C. Pacchierotti, and M. Babel, "Power wheelchair navigation assistance using wearable vibrotactile haptics," *IEEE Trans. Haptics*, vol. 13, no. 1, pp. 52–58, January–March 2020.

[5] J. Vogel, D. Leidner, A. Hagengruber, M. Panzirsch, B. Bäuml, M. Denninger, U. Hillenbrand, L. Suchenwirth, P. Schmaus, M. Sewtz *et al.*, "An ecosystem for heterogeneous robotic assistants in caregiving," *IEEE Robotics & Automation Magazine*, vol. 28, no. 3, pp. 12–28, September 2021.

[6] Age UK, "Later life in the UK factsheet 2019," accessed on April 4, 2022. [Online]. Available: https://www.ageuk.org.uk/globalassets/age-uk/documents/reports-and-publications/later_life_uk_factsheet.pdf

[7] National Health Service (NHS), "Improving care for older people," accessed on April 4, 2022. [Online]. Available: https://www.england.nhs.uk/ourwork/clinical-policy/older-people/improving-care-for-older-people/

[8] M. Sharifi, S. Behzadipour, H. Salarieh, and M. Tavakoli, "Assist-as-needed policy for movement therapy using telerobotics-mediated therapist supervision," *Control Engineering Practice*, vol. 101, p. 104481, 2020.

[9] D. A. Abbink, T. Carlson, M. Mulder, J. C. de Winter, F. Aminravan, T. L. Gibo, and E. R. Boer, "A topology of shared control systems – finding common ground in diversity," *IEEE Transactions on Human-Machine Systems*, vol. 48, no. 5, pp. 509–525, 2018.

[10] C. Ezeh, P. Trautman, L. Devigne, V. Bureau, M. Babel, and T. Carlson, "Probabilistic vs linear blending approaches to shared control for wheelchair driving," in *2017 Int. Conf. on Rehabilitation Robotics (ICORR)*, Toronto, Canada, July 2017, pp. 835–840.

[11] L. Devigne, V. K. Narayanan, F. Pasteau, and M. Babel, "Low complex sensor-based shared control for power wheelchair navigation," in *2016 IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS)*, Daejeon, South Korea, October 2016, pp. 5434–5439.

[12] V. Narayanan, A. Spalanzani, and M. Babel, "A semi-autonomous framework for human-aware and user intention driven wheelchair mobility assistance," in *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS)*, Daejeon, South Korea, October 2016, pp. 4700–4707.

[13] D. Fox, W. Burgard, and S. Thrun, "The dynamic window approach to collision avoidance," *IEEE Robotics & Automation Magazine*, vol. 4, no. 1, pp. 23–33, 1997.

[14] P. D. Nisbet, "Who's intelligent? wheelchair, driver or both?" in *Proc. Int. Conf. on Control Applications*. Padova, Italy: IEEE, September 2002, pp. 760–765.

[15] D. Bertsekas, *Dynamic programming and optimal control*, 3rd ed. Belmont, MA: Athena Scientific, 2005, vol. 1.

[16] C. S. Teodorescu, S. Vandenplas, B. Depraetere, K. Shariatmadar, T. Vyncke, J. Duflou, and A. Nowé, "An ECMS-based powertrain control of a parallel hybrid electric forklift," in *21st Int. Conf. on System Theory, Control and Computing*, Sinaia, Romania, October 2017, pp. 763–770.

[17] S. Liu, L. Li, J. Tang, S. Wu, and J.-L. Gaudiot, *Creating autonomous vehicle systems*, 2nd ed., ser. Synthesis Lectures on Computer Science. Morgan & Claypool, 2020.

[18] Y. Pyo, H. Cho, R. Jung, and T. Lim, *ROS Robot Programming*. ROBOTIS, 2017.

[19] L. Johannesson, M. Asbogard, and B. Egardt, "Assessing the potential of predictive control for hybrid vehicle powertrains using stochastic dynamic programming," *IEEE Trans. Intell. Transp. Syst*, vol. 8, no. 1, pp. 71–83, March 2007.

[20] H. Liu, R. Xu, L. Han, and S. Xiong, "Control strategy for an electromechanical transmission vehicle based on a double Markov process," *Int. Journal of Automotive Technology*, vol. 22, no. 3, pp. 761–770, 2021.

[21] C.-C. Lin, H. Peng, and J. W. Grizzle, "A stochastic control strategy for hybrid electric vehicles," in *American Control Conf.*, Boston, MA, June 30–July 2 2004, pp. 4710–4715.

[22] G. Feng, T. M. Guerra, L. Busoniu, A.-T. Nguyen, and S. Mohammad, "Robust observer-based tracking control under actuator constraints for power-assisted wheelchairs," *Control Engineering Practice*, vol. 109, p. 104716, 2021.

[23] C. S. Teodorescu, B. Zhang, and T. Carlson, "A stochastic control strategy for safely driving a powered wheelchair," in *21st IFAC World Congress*, Berlin, Germany, July 2020.

[24] C. S. Teodorescu, I. Caplan, H. Eberle, and T. Carlson, "Model-based sensor fusion and filtering for localization of a semi-autonomous robotic vehicle," in *European Control Conference*, Rotterdam, The Netherlands, June 29–July 2 2021.

[25] F. Bonnans and P. Rouchon, *Commande et Optimisation de Systèmes Dynamiques*. Palaiseau, France: Èditions de l'Ècole Polytechnique, 2005.

[26] International Automobile Federation, "Formula 1 - Sporting Regulations 2021," accessed on April 4, 2022. [Online]. Available: https://www.fia.com/regulation/category/110

[27] A. Petrie and C. Sabin, *Medical Statistics at a Glance*, 4th ed. John Wiley & Sons, 2019.

[28] C. S. Teodorescu, B. Zhang, and T. Carlson, "A rule-based assistive control algorithm for safe navigation for a powered wheelchair," in *European Control Conference*, Saint Petersburg, Russia, May 2020, pp. 1204–1209.

[29] B. W. Brown and M. Hollander, *Statistics: A Biomedical Introduction*. John Wiley & Sons, 1977.

[30] D. Navarro and D. Foxcroft, *Learning statistics with jamovi: a tutorial for psychology students and other beginners*, 2019.

[31] W. K. B. Hofstee, J. M. F. Ten Berge, and A. A. J. Hendriks, "How to score questionnaires," *Personality and Individual Differences*, vol. 25, no. 5, pp. 897–909, 1998.

[32] P. Trautman, "Assistive planning in complex, dynamic environments: a probabilistic approach," in *2015 IEEE Int. Conf. on Systems, Man, and Cybernetics*, Hong Kong, October 2015, pp. 3072–3078.

[33] B. Zhang, J. Amirian, H. Eberle, J. Pettré, C. Holloway, and T. Carlson, "From HRI to CRI: Crowd robot interaction - understanding the effect of robots on crowd motion," *Int. Journal of Social Robotics*, pp. 1–13, 2021.