

An ‘Ethical Black Box’, Learning From Disagreement in Shared Control Systems

Henry Eberle*, Bingqing Zhang[†], Catalin Stefan Teodorescu*, George Walker*, Tom Carlson*

**Aspire Centre for Rehabilitation Engineering and Assistive Technology, University College London*

[†]*Department of Computer Science, University College London*

Abstract—Shared control, where a human user cooperates with an algorithm to operate a device, has the potential to greatly expand access to powered mobility, but also raises unique ethical challenges. A shared-control wheelchair may perform actions that do not reflect its user’s intent in order to protect their safety, causing frustration or distrust in the process. Unlike physical accidents there is currently no framework for investigating or adjudicating these events, leading to a reduced capability to improve the shared control algorithm’s user experience. In this paper we suggest a system based on the idea of an ‘ethical black box’ that records the sensor context of sub-critical disagreements and collision risks in order to allow human investigators to examine them in retrospect and assess whether the algorithm has taken control from the user without justification.

Index Terms—Shared control, Powered wheelchair, Assistive technology, Safety

I. INTRODUCTION

Powered wheelchairs are an invaluable tool for allowing the physically disabled to navigate a world full of exclusionary design choices, but currently they are of limited benefit to those with cognitive or sensory impairments. Because of the high level of mobility they provide, prescribing authorities are reluctant to provide powered wheelchairs to patients they believe cannot use them safely. Unfortunately, cognitive and sensory impairments are often comorbid with the type of physical impairments these devices were designed to address, leaving a large population of potential users unserved.

Shared control is a paradigm that attempts to address this problem by augmenting the wheelchair user’s control capabilities through blending their commands with those of an algorithmic controller. This differs from *autonomous* control where the user indicates a destination and the wheelchair navigates there itself. Instead, the user performs all those control functions that they are able to, while the controller handles the remainder. This keeps the user engaged with the control of the wheelchair, combating the skill decay that can occur when they remain a passive passenger for long periods.

From a medical and care perspective, the primary objective of a shared control powered wheelchair is the safety of its occupant and those around them. However, it is important to balance this with the user’s psychological need for

self-determination. Potential users have revealed significant concern over loss of control to the wheelchair’s algorithm [1], particularly when the reasons why are opaque to them. Thus there are two primary failure modes for a shared control powered wheelchair: the physical risk of entering a hazardous situation (or failing to prevent the user from doing so), and the ‘ethical risk’ of overriding the user’s valid intention to perform a particular action. A high frequency of the former will affect the wheelchair’s ability to be certified as a medical device, while the latter will adversely affect user experience, potentially reducing engagement. On a deeper level, interfering too much with the user’s control cuts against the central purpose of an assistive device such as a wheelchair: providing the means for the user to realise their intent despite their impairment(s). At its worst, a highly ‘disagreeable’ shared controller could exacerbate rather than relieve its user’s disability.

One suggested method of adjudicating the failures of an autonomous system is through the use of an ‘ethical black box’ [2] [3], an analogue of an aeroplane flight recorder that logs sensor data and other information pertinent to a robot’s decision-making processes for later examination. In the original sense, a vehicle’s recorder is designed to record continually on a buffer until a catastrophic event (usually a crash), allowing the circumstances to be investigated after its retrieval. However, for a shared control wheelchair a failure state may not correspond to an unambiguous crash or disaster. In disordered pedestrian environments hazards and points of disagreement with the user are likely to be transient, and it is desirable for the wheelchair not to ‘freeze’ as this could put the user (and surrounding pedestrians) at further risk. Therefore it is necessary to design a trigger for logging an event to long-term storage that recognises these problematic events without relying on the clear signal of mechanical failure. Highlighting and recording times where the user and algorithm disagreed or the wheelchair manoeuvred into a high-risk situation would allow for post-hoc examination of these events without combing through hours of irrelevant data or requiring expensive storage.

For the purposes of this paper, we address “input-mixing” shared control as defined in the review by Abbink et al. [4]: control systems that receive data from both the driver’s input device(s) and sensors mounted on the wheelchair and

*This project has received funding from the European Union’s Horizon 2020 research and innovation programme under the grant agreement CROWD-BOT No 779942.

combine them algorithmically in order to create the final motor command. This does not encompass autonomous wheelchairs where the user indicates a target location and the wheelchair performs 100% of the navigation, or “traded control” where the human and the algorithm are each in full control of the wheelchair at different points in time. Since there are many disparate approaches to input-mixing shared control in wheelchairs, some of which do not make their internal logic available, we first summarise key state-of-the-art methods in terms of their behavioural properties in section II before defining applicable ‘failure events’ that can be detected algorithmically in order to trigger memorisation in section III. In section IV, we test our recording methodology in a simulated crowd scenario, before discussing how the autonomous recorder can be used in concert with a human adjudicator to identify faults in the shared control algorithm (section V) and outlining our conclusions in section VI.

II. MODERN APPROACHES TO WHEELCHAIR SHARED CONTROL

Input mixing shared control algorithms differ in two primary ways: how they generate their ‘optimal’ trajectory, and how they combine this with the user’s input (the blending method). The two most prominent blending methods are linear blending, where the wheelchair follows a weighted mixture between the user’s command and an ‘optimal trajectory’, and Probabilistic Shared Control (PSC) which treats the ‘preferred’ trajectories of the user and the algorithm as samples of a probability distribution, selecting the trajectory with the highest probability of satisfying both parties [5].

One of the most commonly iterated-upon approaches is the potential field method, which calculates the desired motion of the wheelchair as if it is a passive object acted upon by attractive and repulsive forces. In the most simple, autonomous, form, detected obstacles in the environment repel the wheelchair while the target destination acts as a gravity well, drawing it in [6]. Shared control can be implemented by treating the user’s input as another force, and its priority can be changed by modulating its strength [7]. If the user’s ‘force’ is arbitrarily strong the user will experience full control, while if it is weak it will not overcome the repulsive force of obstacles and the wheelchair will avoid them even when instructed not to. Because the blending method is a linear sum based on (abstracted) Newtonian mechanics, force field algorithms can become ‘trapped’ in local minima of the virtual force field [8]. For this reason, modern implementations incorporate numerous performance-improving modifications such as those of the Vector Field Histogram (VFH) [9] which searches for minima in the distribution of obstacles in a 1-dimensional polar histogram centred on the wheelchair. This prevents many of the effects associated with local minima in a 2-dimensional virtual force field such as oscillations in confined areas and swerving away from narrow entryways. The linear blending of user and algorithmic commands can also be enhanced by contextually

altering their relative weighting, for example giving the user more control when obstacles are scarce [10].

Methods based on the Dynamic Window Approach (DWA) [11] use a different blending strategy, in that a set of physically achievable candidate velocities are created by the motion planner, with one being selected based on an objective function. The goals of a DWA motion planner often include maximising speed and clearance from obstacles while minimising deviations from the desired heading, but the fidelity to the user’s commands can be added as an additional weighted parameter, implementing shared control [12]. The key advantage of DWA is that each trajectory is physically plausible, and is generated such that the wheelchair is capable of stopping before impacting any (detected) obstacle. Probabilistic DWA shared control is an extension of this method that treats the DWA candidates as a discretised probability distribution of trajectories, with their likelihood determined by how well they fulfil the motion planner’s objective function. The user’s input is similarly treated as the centre of a distribution of possible commands, acknowledging that some conditions can reduce the precision of motor function and/or cognitive planning. The algorithm then seeks to output the candidate that maximises the joint probability of these two distributions. Probabilistic DWA can be further extended by applying additional constraints to the candidate velocities: in work by Zhang et al. [13] a hierarchical controller based on DWA uses the Generalized Velocity Obstacle (GVO) method [14] to filter out candidates that are likely to result in a collision with a pedestrian(s).

A less conventional approach, Stochastic Dynamic Programming (SDP) is a method for calculating the optimum control policy under conditions of uncertainty [15]. This is strongly relevant to the shared control case in that the user’s intent can be characterised using a probability distribution and incorporated into this calculation. SDP pre-computes an optimal policy offline satisfying the joint goals of protecting the user’s safety and respecting their desires. Due to the ‘curse of dimensionality’, computing the optimal solution to more complex SDP problems is prohibitively expensive, which has motivated the creation of simpler approximations such as decoupling the forward and lateral motion of the wheelchair [16].

III. A SHARED CONTROL ‘BLACK BOX’

The key shared characteristic of all these shared control algorithms is that they perform a modification on the user’s input, producing behaviours that the user may not recognise as being in line with their intentions. If the user disagrees strongly with the decisions the algorithm makes, they may become justifiably frustrated and refuse to engage with the wheelchair. At the same time users, particularly those with cognitive or communication impairments, may find it difficult to explain the context of the disagreement and thus struggle to build an appropriate mental model.

Many proposed forms of shared control attempt to explicitly model the capabilities and intent of the user, assisting to match their desires and in proportion to their need. However, user skill development, adaptation to the wheelchair, and condition progression can change user requirements over time. While an automated system can track the changing profile of user-wheelchair interaction, the subjective component of disagreement means that a human is better equipped to understand the user’s evolving needs and desires.

Because of the many forms shared control algorithms can take, a broadly-applicable recording function cannot be reliant on accessing specific elements of their internal logic. In addition to the technical difficulties of doing so, developers may be reluctant (or unable) to disclose sensitive details of their algorithm and thus refuse to incorporate such a system into their platform. To accommodate this, the recorder should rely on the available output of the wheelchair’s sensors and define its own metrics for safety and disagreement (this also limits exposure to biases that could be present in the shared-controller’s own performance estimates). For the purposes of a ‘full’ retrospective safety assessment we will assume that the wheelchair is equipped with odometry and at least one form of rangefinding such as lidar or ultrasonic range sensors, and a camera. The latter is assumed to capture a sufficient field of view that the context of recorded events can be examined by a human investigator, potentially with the assistance of structured feedback from the user and/or their carer.

The collection of video data is obviously sensitive, and raises legal issues related to data protection. As per the General Data Protection Regulation (GDPR) and UK-GDPR [17], non-anonymised personal data must not be collected without the explicit permission of the subject, which may be impossible to obtain in a dynamic environment like a crowd. Therefore, any sensor data that could allow the identification of individuals must be anonymised (for example, by automated face blurring [18]) before being transferred from the temporary buffer to permanent storage (i.e. before it is processed). Although this may destroy some information, anonymised video is still of higher value for determining the context of an event than other, lower-resolution sensor systems such as ultrasonic rangefinders.

An existing set of metrics for observational shared-control experiments was described by Zhang et al [13]. Of these, the most important is ‘agreement/disagreement’ which is defined by the magnitude of the difference between the user’s command and the final movement of the wheelchair. Notably, this does not require accessing the algorithm’s ‘preferred’ trajectory, only the final movement as recorded by the sensors. The disagreement between user and controller is here defined as

$$d = \|u - v\| \quad (1)$$

where u is the instantaneous velocity command from the user and v is the instantaneous velocity of the wheelchair (the use of velocity here is motivated by the fact that most shared-control wheelchairs are built on top of powered wheelchair models that natively use velocity control). Because some input devices may output at a lower frequency than the wheelchair controller, the input and output velocities are held in a zero-order-and-hold buffer and the disagreement is calculated from the last recorded signals. This leads to natural variations in the agreement that do not necessarily indicate an error in the shared control. For example, a head array may output a highly discretised, low frequency signal that will naturally differ significantly from the continuous motion of the wheelchair as it moves through a crowd. For this reason, a disagreement threshold T_d is defined as

$$T_d = AD \quad (2)$$

where D is the expectation of disagreement, defined as the lowest agreement that would be expected under optimal conditions, and A is a constant of sufficient size to exclude deviations due to sensor error. The trigger to begin recording is defined as

$$T_d < \tilde{d}_t = \frac{\sum_{i=t-\tau}^t d_i}{T/\Delta t} \quad (3)$$

where \tilde{d}_t is the average disagreement over the previous τ seconds and Δt is the inverse of the recorder’s sampling frequency. Upon average disagreement exceeding T_d , a marker is created for n seconds before the event to indicate the beginning of record, and a second marker is created n seconds after the average disagreement has dropped below T_d . The sensor and input device output between these markers is then consigned to long-term storage.

Because many instrumented wheelchairs do not have true contact sensing, and because range sensors can become unreliable at very close range, we identify scenarios with a high risk of a collision by estimating whether the wheelchair could decelerate to a stop in the time (t_h) before impacting the nearest obstacle, defining a hazard Boolean value, h :

$$t_h = v/a \quad (4)$$

$$h = \begin{cases} 1, C < st_h - 0.5at_h^2 + b \\ 0, C \geq st_h - 0.5at_h^2 + b \end{cases} \quad (5)$$

where C is the distance to the obstacle in metres, s is the instantaneous speed of the wheelchair, and a is the wheelchair’s maximum acceleration opposite to the current direction of travel. If the wheelchair is close enough to an obstacle that it will reach it within the t_h seconds required to halt itself a high risk of collision can be inferred even without contact sensors. The constant b represents uncertainty in the relative velocity between wheelchair and a potentially moving obstacle, the detection of which can be unreliable using low-frequency range sensors. A higher b results in a

more ‘cautious’ algorithm, recording events with a lower likelihood of collision. To reduce the influence of sensor errors the instantaneous hazard value h is also averaged over τ_s , with the hazard condition only triggering if at least 50% of the last $\tau/\Delta t$ samples are equal to 1.

The two error scenarios (risk and disagreement) are independent, so each recorded time step is further given a Boolean tag for each to indicate why it has been committed to memory. The events requiring examination can thus be divided into: disagreement leading to hazard, disagreement leading to no hazard, and agreement leading to hazard (agreement leading to no hazard is the normal, desired behaviour). While the first and last of these appear to be clear failures of the algorithm, overriding the user to make a risky manoeuvre and erroneously agreeing with the user’s risky actions, respectively, all will ultimately require human examination. In every case, there is the potential for contextual information that could justify the driver’s ‘erroneous’ input, and apparent hazards can be the result of sensor artifacts.

Thus all the above automated triggers are designed to work in concert with a human investigator, who will examine the recorded events and determine whether the shared control algorithm made an error, and of what severity. Because a trained human is much more likely than an automated system to be able to adjudicate the difference between a situation that required the algorithm to take control and one the user could handle themselves, the recorder system should be designed to err on the side of logging ambiguous situations.

IV. CASE STUDY IN A SIMULATED ENVIRONMENT

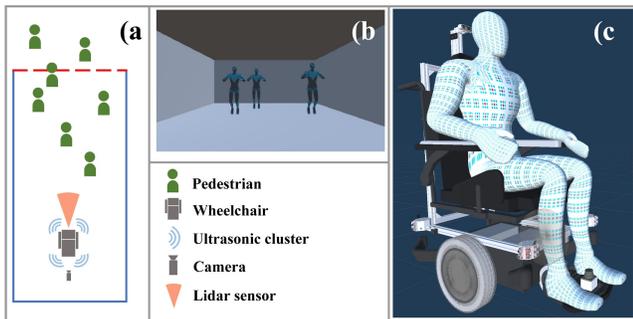


Fig. 1. Layout (a) of the simulated crowd scenario (truncated for display purposes), a wheelchair-perspective view (b) of the corridor within Unity, and the simulated wheelchair itself (c). The pedestrians and the wheelchair travel in opposite directions through the corridor. The far end of the corridor (dashed red in (a)) allows pedestrians through while blocking the wheelchair, and marks the end of a run when reached by the wheelchair.

In order to assess the applicability of this system to classical pedestrian scenarios, a crowd scene (Fig.1 (a)) was simulated using Unity and designed with a high potential for collision and disagreement events. A simulated wheelchair (Fig.1 (c)) was tasked with moving down a 15m by 6m corridor (approximately 15x the simulated wheelchair’s length).

Modelled on our smart wheelchair [15], the simulated wheelchair was equipped with 12 simulated ultrasonic sensors situated in clusters of three on each corner of the wheelchair, as well as a frontally-mounted LIDAR sensor, with a tracking Unity camera acting as a camera sensor. Each ultrasonic sensor had a maximum range of 1.5m and field of view of 45°, while the LIDAR had a maximum range of 5.6m and a 180° field of view, collectively giving the wheelchair 360° sensing. Randomly-placed ‘humans’ moved at pedestrian speed (1.1 m/s) down the corridor in the opposite direction to the wheelchair, presenting a high chance of a hazardous collision. The use of a simulation allowed this scenario to be repeated multiple times with randomised initial positions for the pedestrians, thus providing multiple points at which the algorithm is expected to disagree with the driver and multiple collision events at no risk to the tester.

The wheelchair was controlled in concert with one of two shared control algorithms, with the user input provided by two non-wheelchair users who were instructed to drive the wheelchair from one end of the corridor to the other in as close to a straight line as possible, ignoring pedestrians. Both users had significant experience using the simulated wheelchair in similar experiments. The first algorithm sought to artificially introduce conflict with the user by linearly mixing their input with a random command vector of up to 50% the wheelchair’s maximum velocity, changing every 5s. This generates what Itoh et al. characterised as “intention conflict” [19]: the driver intends to reach the end of the corridor, while the algorithm intends to travel in a random, changing direction.

The second algorithm was a PSC DWA-GVO controller developed in earlier work by Zhang et al. [13] that selects collision-free trajectories that present the highest chance of fulfilling the user’s intentions while guaranteeing safety. As the drivers were instructed to ignore pedestrians, this was expected to lead to both intention conflict (because the algorithm intends to avoid all collisions) and information gathering conflict, as the driver has access to visual information that the algorithm does not (Fig.1 (b)).

The recorder’s ability to capture hazards was assessed by how many of its trigger events came within 1s of a collision vs. how many triggered without a collision (true positives vs. false positives) as well as how many collisions occurred without being flagged, with collisions being defined as intersections between the wheelchair model and either a wall or a pedestrian. The disagreement detection was assessed qualitatively by its ability to distinguish ‘justified’ causes of disagreement (proximity to walls or pedestrians for the PSC DWA-GVO algorithm and ‘swerve’ events for the random algorithm) from the normal operation of the wheelchair (transient differences between commanded and actual velocity due to limited acceleration).

The simulation was repeated 20 times each for the PSC DWA-GVO controller and the random controller, with a single run lasting from the point the wheelchair began moving to when it either reached the far end of the corridor or became immobile due to collision. The initial positions of the 8 moving obstacle ‘pedestrians’ was drawn from a pseudorandom tuple generated by C#’s `random.range()` function.

During the simulation, the recorder’s time step Δt was set to 0.1s, the length of the averaging filter T was set to 0.3s, the disagreement threshold was set at 0.5 (twice the expected average disagreement), and the hazard uncertainty constant b was set to 0.4. The simulated wheelchair’s maximum forward linear acceleration was $0.26m/s^2$, its maximum backwards linear acceleration was $1m/s^2$, its maximum angular acceleration was $0.4rad/s^2$ and its maximum speed was $1m/s^2$. The user command to the simulated wheelchair was provided by a joystick with the vertical and lateral axes mapped to linear and angular velocity, respectively.

A. Results

Over all 20 PSC DWA-GVO runs (Table I), 8 out of 17 collisions with pedestrians triggered recording due to disagreement, and a further 8 were recorded within a longer period of disagreement caused by passing through a high density of pedestrians, leaving only 1 unrecorded. Disagreement was highest when passing through dense groups of pedestrians, as seen in Fig.2 (a), with 37 out of 55 instances of disagreement preceded by the approach of a pedestrian. Of the remainder, 5 were erroneous recordings triggered by the wheelchair’s slow acceleration, while the remainder had no explanation that could be inferred from the recorded data. Only 9 instances of hazard were recorded, none of which preceded a collision, but all of which corresponded to coming to within 0.8m of a pedestrian.

During the 20 random controller runs 19 total collisions were recorded (Table I), with 15 flagged as hazards, 6 flagged as disagreements and 1 recorded during a long period of disagreement (3 collisions were flagged as both disagreement and hazard). Out of the 226 recorded disagreement events, 171 were short ‘steering events’ associated with correcting the wheelchair’s heading (as seen in Fig. 2 (b)), while 55 were either associated with acceleration delays or had no cause that could be inferred from the recorded data. 26 of the remaining 28 false positive hazards were associated with coming within 0.8m of a pedestrian.

V. DISCUSSION

The recorder system demonstrated a consistent ability to record circumstances where the driver disagreed with the movements of the wheelchair, although it showed some difficulty in distinguishing between points where the wheelchair’s motion differed due to intervention by the algorithm, and where it was merely slow in responding to

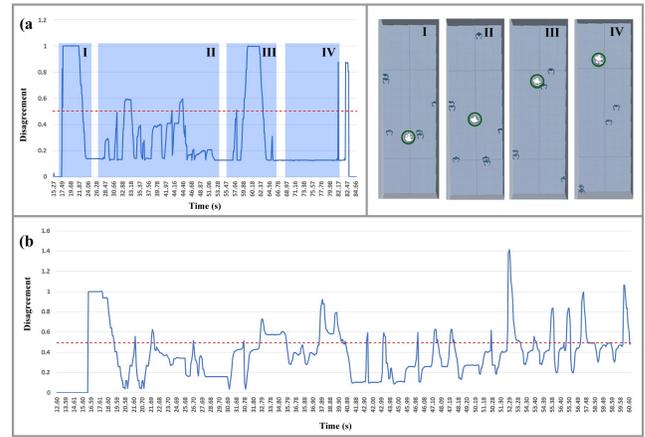


Fig. 2. Disagreement over time (beginning from first movement) for a single run of the simulation using (a) PSC-DWA-GVO and (b) the random controller. While disagreement in (b) occurs in regular ‘spikes’, periods of disagreement in (a) are associated with proximity to pedestrians (inset screenshots I-IV, with wheelchair circled in green).

abrupt changes in input. This reflects that the implementation was based on a ‘worst case scenario’ where only the minimum viable data (the user command and sensor feedback) was available. Despite this, the majority of pertinent hazards were detected. The primary cause of false-positive disagreements could be eliminated by basing the disagreement on the final blended velocity command rather than the velocity recorded by the sensors, which is affected by unpredictable dynamics and flawed numerical differentiation. Future studies should ideally test whether the sensor feedback or the velocity command is a closer match to the user’s subjective experience, as disagreement is ultimately supposed to represent divergence from the user’s intentions.

Since the PSC DWA-GVO algorithm either swerved around pedestrians or brought the wheelchair to a stop (which would allow them to walk around in a real scenario), all collisions occurred at very low velocities that would be safe in practice. As such, all instances of the hazard trigger for the PSC DWA-GVO were false positives caused by a pedestrian passing close to the wheelchair without actually impacting it. This was partially due to a mismatch between how the simulated wheelchair’s sensors reported the distance to the nearest obstacle and how the Unity engine registered collisions, but implies that any future version of the recorder should incorporate an estimate of oncoming obstacles’ shape in order to avoid recording such ‘near misses’.

Tests using the random controller registered multiple collisions, the majority of which were detected by the recorder’s hazard condition, usually after the collision itself. Thus in terms of the framework discussed in Section. III, the recorder mostly flagged cases of “disagreement leading to no collision” for the PSC DWA-GVO, and “agreement leading to collision” for the random controller. In most cases

TABLE I
SIMULATION RESULTS

	Collisions				Disagreements		Hazards	
	Undetected	Hazard Flag	Disagreement Flag	Recorded (No Flag)	Justified	Unjustified	Collision	No Collision
PSC-DWA-GVO	1	0	8	8	37	18	0	9
Random Controller	0	15	6	1	171	55	15	28

the PSC DWA-GVO algorithm was disagreeing in order to prevent a potential collision, but given that slowing down in a busy area is in and of itself a risk, this is exactly the type of scenario that requires human examination in a real-world scenario. Triggering too late was also the primary cause of ‘missed’ collisions. As these delays extended beyond the length of the averaging filter, this was again most likely caused by a mismatch between the point at which the Unity engine detected a collision and the distance reported by the wheelchair’s sensors.

The majority of ‘steering’ events where the user attempted to counteract the random controller’s changes of direction were flagged by the recorder, although there were also a significant minority of disagreements that did not appear to correlate with any features of the recorded sensor data. This highlights a crucial consideration: the disagreement, as indicated by the recorded data, can emerge significantly after the event that ultimately caused it. The user may exhibit a slow or highly variable response time, or they may only begin disagreeing with the algorithm’s decision once the longer-term consequences become apparent. Thus although the disagreement is recorded, the point at which the algorithm could have feasibly prevented it is not. Enlarging the recording window around trigger events may ameliorate this problem, but it is important that this is balanced against the ability of human examiners to sift through greater amounts of data.

VI. CONCLUSION

Shared control presents challenges beyond simply preserving the safety of a user. Although an algorithm may be excellent at preventing harm, over-intervention can leave a user feeling disempowered, which is at odds with the aims of an assistive device such as a powered wheelchair. Unfortunately, many of the users of shared control wheelchairs suffer from impairments that may make communicating the cause of their frustration difficult. Thus it is important for any ‘ethical black box’ that is applied to the shared control case to not only record physical accidents, but also points at which the user appears to be struggling against the assistive algorithm. This gives investigators the best chance of understanding the context of these disagreements, and can provide valuable data for the improvement of the user experience.

REFERENCES

- [1] P. Viswanathan, E. P. Zambalde, G. Foley, J. L. Graham, R. H. Wang, B. Adhikari, A. K. Mackworth, A. Mihailidis, W. C. Miller, and I. M. Mitchell, “Intelligent wheelchair control strategies for older adults with cognitive impairment: User attitudes, needs, and preferences,” *Autonomous Robots*, vol. 41, no. 3, pp. 539–554, 2017.
- [2] A. F. Winfield and M. Jirotko, “The case for an ethical black box,” in *Annual Conference Towards Autonomous Robotic Systems*. Springer, 2017, pp. 262–273.
- [3] A. F. Winfield, K. Winkle, H. Webb, U. Lyngs, M. Jirotko, and C. Macrae, “Robot accident investigation: a case study in responsible robotics,” *arXiv preprint arXiv:2005.07474*, 2020.
- [4] 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.
- [5] P. Trautman, “Assistive planning in complex, dynamic environments: a probabilistic approach,” in *2015 IEEE International Conference on Systems, Man, and Cybernetics*. IEEE, 2015, pp. 3072–3078.
- [6] M. R. Petry, A. P. Moreira, R. A. Braga, and L. P. Reis, “Shared control for obstacle avoidance in intelligent wheelchairs,” in *2010 IEEE Conference on Robotics, Automation and Mechatronics*. IEEE, 2010, pp. 182–187.
- [7] C. Urdiales, J. M. Peula, M. Fdez-Carmona, C. Barrué, E. J. Pérez, I. Sánchez-Tato, J. Del Toro, F. Galluppi, U. Cortés, R. Annichiarico *et al.*, “A new multi-criteria optimization strategy for shared control in wheelchair assisted navigation,” *Autonomous Robots*, vol. 30, no. 2, pp. 179–197, 2011.
- [8] Y. Koren, J. Borenstein *et al.*, “Potential field methods and their inherent limitations for mobile robot navigation,” in *ICRA*, vol. 2, 1991, pp. 1398–1404.
- [9] J. Borenstein and Y. Koren, “The vector field histogram-fast obstacle avoidance for mobile robots,” *IEEE Transactions on Robotics and Automation*, vol. 7, no. 3, pp. 278–288, 1991.
- [10] Qinan Li, Weidong Chen, and Jingchuan Wang, “Dynamic shared control for human-wheelchair cooperation,” in *2011 IEEE International Conference on Robotics and Automation*, 2011, pp. 4278–4283.
- [11] 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.
- [12] P. Inigo-Blasco, F. Diaz-del Rio, S. V. Diaz, and D. C. Muniz, “The shared control dynamic window approach for non-holonomic semi-autonomous robots,” in *ISR/Robotik 2014; 41st International Symposium on Robotics*. VDE, 2014, pp. 1–6.
- [13] B. Zhang, C. Holloway, and T. Carlson, “A hierarchical design for shared-control wheelchair navigation in dynamic environments,” in *2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. IEEE, 2020, pp. 4439–4446.
- [14] D. Wilkie, J. Van Den Berg, and D. Manocha, “Generalized velocity obstacles,” in *2009 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 2009, pp. 5573–5578.
- [15] C. S. Teodorescu, B. Zhang, and T. Carlson, “A stochastic control strategy for safely driving a powered wheelchair,” in *2020 IFAC World Congress*, 2020.
- [16] —, “A rule-based assistive control algorithm for safe navigation for a powered wheelchair,” in *2020 European Control Conference (ECC)*. IEEE, 2020, pp. 1204–1209.
- [17] Council of European Union, “Regulation (eu) 2016/679 of the european parliament and of the council of 27 april 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing directive 95/46/ec (general data protection regulation),” 2016.
- [18] E. Flouty, O. Zisimopoulos, and D. Stoyanov, “Faceoff: anonymizing videos in the operating rooms,” in *OR 2.0 Context-Aware Operating Theaters, Computer Assisted Robotic Endoscopy, Clinical Image-Based Procedures, and Skin Image Analysis*. Springer, 2018, pp. 30–38.
- [19] M. Itoh, F. Flemisch, and D. Abbink, “A hierarchical framework to analyze shared control conflicts between human and machine,” *IFAC-PapersOnLine*, vol. 49, no. 19, pp. 96–101, 2016.