This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/LRA.2021.3136646, IEEE Robotics and Automation Letters

The Biasing of Action Selection Produces Emergent Human-Robot Interactions in Autonomous Driving

Mauro Da Lio[1], Riccardo Donà[1], Gastone Pietro Rosati Papini[1], Alice Plebe[1]

***Abstract*—This paper describes a means to produce emergent collaboration between a human driver and an artificial co-driver agent. The work exploits the hypothesis that human-human cooperation emerges from a shared understanding of the given context's affordances and emulates the same principle: the observation of one agent's behavior steers another agent's decision-making by favoring the selection of the goals that would produce the observed activity. Specifically, we describe how to steer the decision-making of a special self-driving agent via weighting the agent's action selection process with input from a dummy human driving activity. In this way, human input maps onto the safe and affordable actions recognized by the agent. We demonstrate an emergent and efficient driving, collaboration, and rejection of unsafe human requests.**

***Index Terms*—Cognitive Modeling; Human Factors and Human-in-the-Loop; Human-Robot Collaboration**

I. INTRODUCTION

THE rider-horse metaphor (H-metaphor) was proposed back in 2003 as a model of desirable interaction between drivers and intelligent vehicles [1]. Norman [2, pag. 19] outlines the salient aspects of the interaction as follows:

> *Think of skilled horseback riders. The rider "reads" the horse, just as the horse can read its rider. (...) This interaction is (...) of special interest because it is an example of two sentient systems, horse and rider, both intelligent, both interpreting the world and communicating their interpretations to each other.*

We argue that autonomous vehicles might benefit from a similar ability: the user experience would improve if the driver could give hints to the car and feel as if the car could "understand" his/her intentions.

While the H-metaphor describes a desirable form of interaction, it does not tell how to realize a system that produces that interaction. One way to implement interactions between humans and robots is by programming interplays with rules.

Manuscript received: September, 7, 2021; Revised: November, 2, 2021; Accepted: December, 13, 2021.

This paper was recommended for publication by Editor Tetsuya Ogata upon evaluation of the Associate Editor and Reviewers' comments.

This work was supported in part by the European Union (EU) Horizon 2020 Dreams4Cars Research and Innovation Action with the European Commission Grant 731593. The work was also supported in part by the University of Trento with project Digital Vehicles on Digital Roads (VDSD).

This work is part of a study approved by the ethical committee of the University of Trento (prot. 2018-024).

[1]M. Da Lio, R. Donà, G.P. Rosati Papini, and A. Plebe are with the Department of Industrial Engineering, University of Trento, Trento, Italy. {mauro.dalio, riccardo.dona, gastone.rosatipapini, alice.plebe}@unitn.it

Digital Object Identifier (DOI): see top of this page.

However, an alternative approach may be to construct the robot's sensorimotor system so that the interactions are *emergent*, as much as possible. This paper follows the latter theory.

A. What this paper is (and is not) about

The paper is about replicating (functionally) a natural cognition process that permits an intelligent agent to understand the intentions of another agent and comply with them if possible. The process, also known as "mirroring" (albeit this term may look abused here), has been known for some time, but it is not completely clear how to make it work with an artificial agent. This paper studies a mechanism for that (the biasing of the agent's action selection). To show this mechanism in operation, we need a human to carry out actions that the artificial driver agent interprets. However, the work *is not* about human factors and ergonomy (not about how humans feel and react). The focus is on influencing the agent, which (among others) justifies using a simplified input device. In addition, the sensorimotor system of self-driving agents must have particular characteristics, which the paper describes. Hence, any compatible self-driving agent could work. For practical reasons, we reuse the self-driving agent of the Dreams4Cars project. In this work, we describe the novel interaction mechanics and the enabling elements (Section III), *but no*t a comprehensive description of the rest of the agent, which was published in [3] (agent architecture), [4] (offline learning via mental simulations) and [5] (learning cautious behaviors).

B. Paper organization and contribution

The paper is organized as follows. Section II presents the related literature. Section III describes the self-driving agent characteristics that are necessary for the interaction. Section IV introduces the principles that give birth to emergent interactions. These are grounded in natural cognition theories, such as the simulation hypothesis and the affordance competition hypothesis (Section II). Section V presents an evaluation of the interaction paradigm in driving simulations. We show how various emulated human driver actions can influence the co-driver agent behavior. We also show that the co-driver interprets the human driver hints in terms of what is affordable in the current situation – producing a logical behavior – and can even veto some requests if they are considered dangerous. Section VI deals with generalizations to broader scenarios and discusses a complementary interaction modality that might be necessary for some circumstances.

II. RELATED WORK

In the context of highly automated driving, the H-metaphor [1] was studied in the EU FP7 HAVEit project, realizing a variable repartition of authority between a human driver and a co-pilot system. The implementation uses “H-mode” [6], a “language” for describing and programming human-system interactions. For example, [7] describes four automation levels and presents a case corresponding to driving in a road with traffic alternating car following and changes to free lanes. Readers may find a review of “H-mode” and similar approaches in [8].

Compared to those studies, this work uses a different —genuinely “artificial cognitive”— approach, which relies on the re-use of the agent’s motor system to align two agent’s intentions and produce emergent collective behaviors.

Human-robot interaction modeled on natural human-human collaboration is a vast and active research field, broadly based on the mutual inference of intention.

There are two approaches to infer intentions (see also discussion in [9, Section IV.A]). One utilizes classification: particular patterns in one agent’s sequence of actions —either a human or a robot— are interpreted as the prelude to the next action. It is called *action-to-goal* in [10]. The opposite approach, or *goal-to-action* [10], assumes to deal with intentional agents; i.e., agents whose actions are explainable as directed to the achievement of one goal. By first detecting the attainable goals in a given environment, observers can “imagine” the actions they would use for each goal and then use the observations to accrue evidence of which might be the observed agent’s most likely intention.

The main issue with the *action-to-goal* approach is that action patterns depend on the context: the same pattern may mean different intentions in (even slightly) different contexts, and the same intention may be enacted with varying patterns in other contexts. To clarify, let us make an example of a human driver setting the right indicator light. This action may mean different things in different environments: in a motorway, it may mean the driver wants to stop on the emergency lane (if driving on the rightmost lane with an emergency lane on the right); in the proximity of an intersection, the same action may mean the driver wants to turn right; lastly, on one road with vehicles parked on the right side of the lane, the same action may mean “seeking for a free spot”, which also means that the car will almost stop in the lane before maneuvering to the found spot. The classification approach works within given environments and would require different classifiers for different environments.

On the other hand, for the *goal-to-action* approach to work, the sensorimotor systems of the two agents must be compatible [11]; i.e., the co-driver must be capable of seeing the action possibilities latent in the environment (affordances, [12]) and generating the corresponding action plans in a way very similar to a human being, as a horse sees the affordable paths like humans. For a self-driving agent with only locomotion capabilities, the affordances are navigable corridors, with space-time restrictions from moving obstacles. We further distinguish between physically traversable spaces (e.g., the whole road) and legal corridors of the traffic rules and assign them a different role in a hierarchical sensorimotor system.

Notably, the inference of intentions between human beings happens with the reuse of the sensorimotor system [11], [13], [14]. While observing another person’s activity, the observer’s motor system is engaged —without actual execution of the action— in a way similar to what would be necessary to produce the same action [15]. Hence, the mirroring process estimates the observed agent’s sensorimotor system states that explain the observations.

With the above elucidation, let us restate that this work deals with the *goal-to-action* approach. Researchers have studied possible techniques to reproduce this mechanism for human-robot collaboration. One is the MOSAIC architecture [16], [17], which runs several predictive models in parallel (one per potential goal) to test different hypotheses about the intent of the observed agent. A very similar approach is the HAMMER architecture [18], [19], which tests hypotheses about possible goals versus observations, accumulating evidence (salience) of which one is the most likely. To our best knowledge, the first use of a *goal-to-action* approach in the automotive domain can be found in [9] for a driver assistance system. In that work, potential goals in a driving scenario are resolved into the instantaneous lateral and longitudinal control that a human driver would produce, which are then compared to the actual control used by the driver [9, Fig. 7].

The study of human sensorimotor architectures can be a guide to creating systems that emulate their behavior. In particular, the affordance competition hypothesis [20] describes how potential actions are primed simultaneously in the human “dorsal stream” and encoded in the “motor cortex” with patterns of neural activation, whose level of activation (salience) encodes the expected value of each action. Once the panorama of the possible actions with relevant values is formed, the most salient one may be selected. The selection process in the human brain is carried out in the basal ganglia [21]. It handles noise in the motor system and sensory system by accumulating evidence of which action is likely to be the most salient notwithstanding noise [22], [23]. The action selection process can be biased to pursue longer-term goals, for example, by selecting actions that create the precondition for the long-term goal [24].

III. CO-DRIVER AGENT SENSORIMOTOR SYSTEM

As noted, the sensorimotor architecture of the driving agent must satisfy two requirements. First, it must be compatible with the human (the “like-me” hypothesis [11]), in the sense that the agent recognizes the same affordances that a human would see and creates similar action plans. Second, a mechanism must permit the human to “steer” the driving agent’s choice by indicating which affordance they prefer.

This section deals with the driving agent architecture in itself. The following section deals with steering the agent choice. Because of the interaction with a human, the driving agent is also called “co-driver” [9].

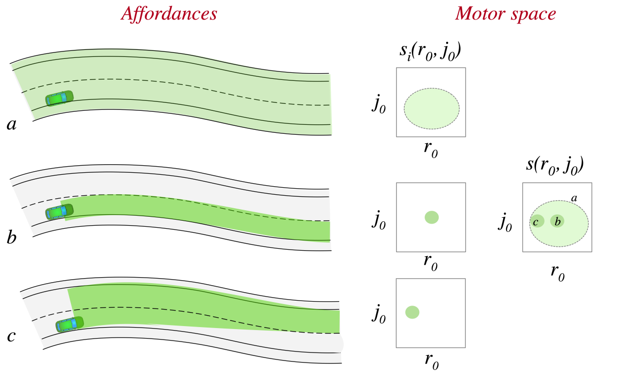

Fig. 1. Action priming. Different affordances are mapped onto active regions on the motor space (center). These are later combined into a general salience function (right) that indicates the optimal choices. By weighting the salience it is possible to steer the agent final choice.

A. Layered control architecture and affordance competition

The co-driver agent (with no interaction) is described in detail in a previous work [3]. We recall here only the elements necessary to enable the interaction mechanism.

The co-driver sensorimotor system is made with a layered control architecture [25]: a cognitive architecture that follows the putative organization of the human sensorimotor system into 1) an *action priming* phase, and 2) a following *action selection* phase. The priming step aims to detect the affordances and "priming" motor plans for all of them simultaneously. The goal is mapping the perceived affordances onto estimates of their "salience" (see next) via perception-action associations. The selection step must choose one affordance, and the choice is carried out with a centralized mechanism based on affordance competition [20].

Layered control architectures have very attractive features. First scalability: if new action possibilities are discovered, the agent needs only to learn the corresponding priming loop, and the newly discovered affordances, encoded in a standard salience scale, become immediately available for competition. Second, it is possible to steer (or bias) the action choice by artificially magnifying or reducing the salience of specific affordances [24]. This mechanism can be exploited internally by the agent itself (e.g., [3, III-C]) or for interaction, which is the purpose of the present work. Finally, adaptive behaviors emerge from the continuous reiteration of affordance competition (e.g., [3, III-A]).

B. Action priming

Priming means the instantiation of the salience of each affordance and its encoding in an appropriate space. The process is carried out with excitatory and inhibitory loops.

1) Excitatory loops: Excitatory loops implement associations between navigable spaces and the corresponding salience. The navigable space is organized in a hierarchy of affordances, as explained in Fig.1. On the top stands the physically traversable space (*a*), e.g., the road including the lateral banks. At the second level stands the legal space, such as, e.g., lanes (*b* and *c*). In Fig.1 there are three affordances: *a*, remain in the road, *b* follow the lane and *c* change lane. The first corresponds to physical integrity, and the latter two are legal requirements. Note that if the centerline were continuous, *c* would not exist.

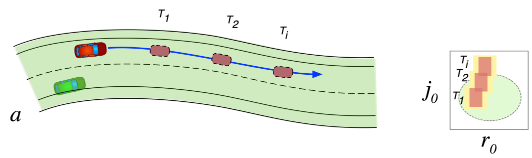

Fig. 2. Action inhibition. Future space-time locations are mapped onto exclusion regions in the motor space.

The actions that the agent may produce are trajectories that originate from the current configuration. Since a vehicle has two controllable degrees of freedom, the whole space of possible actions is spanned by the specification of the longitudinal control $j(t)$ and the lateral control $r(t)$[1]. The dynamics of a vehicle are in part stochastic because of external disturbances: an action $u = \{j(t), r(t)\}$ may generate a family of trajectories γ. The stochastic vehicle response $\{j(t), r(t)\} \to \gamma$ is specified by a probabilistic motion model (in our case, probabilistic motion models were learned with a technique similar to [4]). We hence begin with a mapping $u \to \gamma$ from a generic action $u = \{j(t), r(t)\}$ to the distribution of generated trajectories γ. For every family of trajectories γ, a scalar functional $V(u)$ can be defined to represent how good or desirable that action may be. Different affordances i have a different functional $V_i(u)$ because of the different spatial domains in which the trajectories γ must stay.

In our implementation (but other implementations are possible), the functionals $V_i(u)$ have two factors: 1) for the quality of the lateral control, the probability of remaining in the specified spatial domain i ($i \in a, b, c$) for a sufficient time (alternatively the time resting in the domain i given a threshold probability), 2) for the quality of the longitudinal control, the travel time, subject to speed limits and comfort criteria such as the speed in curves [26].

Since the agent is primarily concerned with selecting the *current* control $\{j(0) = j_0,\, r(0) = r_0\}$ —in adaptive behavior future controls can be modified later—, a definition of salience as a means to express how good the choice of $\{j_0,\, r_0\}$ may be in relation to affordance i, can be given as follows:

$$s_i\,(r_0, j_0) = \sup_u \left(V_i(u) | r(0) = r_0,\, j(0) = j_0\right). \qquad (1)$$

This means that the salience of the instantaneous choice $\{j_0,\, r_0\}$ for affordance i is the value $V_i(\tilde{u})$ of the optimal action $\tilde{u}$ among all actions beginning with $\{j(0) = j_0,\, r(0) = r_0\}$. In Fig.1 the boxes in the central column represent the contour plots of the salience $s_i(r_0, j_0)$[2].

[1] In our implementation $j(t)$ is the longitudinal jerk and $r(t)$ is the steering rate. However, any equivalent set of input could work.

[2] It is not difficult to recognize the similarity with reinforcement learning, where $s_i(r_0, j_0)$ is the Q function estimating the future reward for choosing action $\{j_0,\, r_0\}$ (we have one Q function per affordance).

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/LRA.2021.3136646, IEEE Robotics and Automation Letters

A global salience function $s(r_0, j_0)$ can then be obtained via aggregation of the individual ones, as in Fig.1, right. One possible aggregating function is a weighted $\max$ operator:

$$s(r_0, j_0) = \max(w_i s_i(r_0, j_0), i \in \mathit{affordances}). \quad (2)$$

where weights w_i may serve to steer action selection and to prioritize the affordances. For example, $w_a \ll w_{b,c}$ prioritizes the choice of either *b* or *c* over *a* (i.e., the agent will use road banks as a last resource).

2) Inhibitory loops: Inhibitory loops implement associations between prohibited space-time locations and corresponding points of the motor space where the salience must be reduced (partial inhibition) or zeroed (total inhibition). An example is shown in Fig.2. For moving obstacles, the future prohibited space-time locations are separately computed from a precomputed prediction of the objects' trajectories. The prohibited place of fixed obstacles is straightforward.

For every spatial affordable region, e.g., *a* in Fig.1, the prohibited locations at times T_i are embedded into (1), e.g., by excluding trajectories γ intersecting the location. Every location maps onto an exclusion subregion of the motor space (Fig.2, right). An early example of this computation was given in [27].

C. Action Selection

The instantaneous control $\{j_0,\, r_0\}$ is chosen (with robust algorithm [3, Section II-D]) as the argument of the highest salience in the global salience map (Fig.1, right). Once $\{j_0,\, r_0\}$ is given, it is possible to step backward in the priming loop and find which affordance causes the maximum and generate $\{j(t), r(t)\}$.

IV. PRINCIPLE OF INTERACTION

There are two ways for biasing the action selection process. One intervenes in the global salience function. For example, different regions of the motor space can be weighted as follows:

$$s'(r_0, j_0) = s(r_0, j_0) w(r_0, j_0). \quad (3)$$

where $s'(r_0, j_0)$ is a modifyed salience obtained form $s(r_0, j_0)$ with weights $w(r_0, j_0)$. The second option intervenes at the level of the individual affordances, by varying the weights w_i in (2).

A. Longitudinal bias

The longitudinal bias exploits the first mechanism. In (3) the weights $w(r_0, j_0)$ are redefined as follows.

$$s'(r_0, j_0) = s(r_0, j_0)(g - b)kj_0. \quad (4)$$

i.e., $w(r_0, j_0) = (g - b)kj_0$, where *g* and *b* are the gas and brake strokes normalised to full stroke (i.e., reduced in the range 0-1) and *k* is a convenient gain. In this way, when the human driver presses the gas pedal, she favours the choice of faster affordance (weight proportional to the longitudinal control j_0). The opposite happens if she presses the brake.

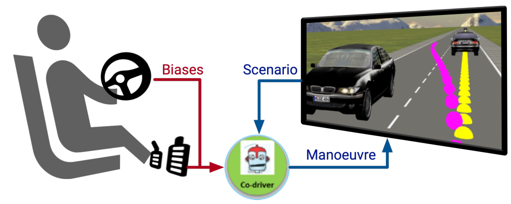

Fig. 3. Schematic representation of the interaction scheme between the human and the artificial driving agent. By acting on dummy actuators, the human driver can bias the decision-making of the "co-driver".

B. Lateral bias

The lateral bias exploits the second mechanism. If we consider the case with both one lane on the right (*R*) and one on the left (*L*), (2) may be rewritten as follows.

$$s(r_0, j_0) = \max(w_i s_i(r_0, j_0), i \in \{L, C, R, a\}). \quad (5)$$

where (*C*) is the current lane and *a* is the drivable road. To influence the choice of the left or right affordable lanes, the weights w_R and w_L (as used by the co-driver in the autonomous mode) are replaced by modified weights w'_R and w'_L, defined as:

$$\begin{aligned} w'_R &= w_R \max(0.1, (1 + \alpha)), \\ w'_L &= w_R \max(0.1, (1 - \alpha)) \end{aligned} \quad (6)$$

where α is the normalized steering wheel rotation (for a conventional rotation that encodes the maximum effect). The maximum lateral bias doubles one weight and reduces at 10% the other.

C. Implementation

In our implementation, the human driver produces the biases α, g and b through *inert* steering wheel and pedal actions (Fig.3). The steering wheel rotation changes the weights of the left and right lane affordances as in (6) and, simultaneously, the actions on the pedals modify the salience according to (4).

Notably, the biasing mechanism works safely. For example, a human's request for faster travel is interpreted in the context of affordable actions —just like the horse maps the rider's wishes into the world. It only favors the choice of faster affordance if it exists (if it does not exist, nothing happens). Similarly, if the human acts on the steering wheel, he/she favors the choice of an affordable lane in that direction if it exists (if it does not exist, nothing happens).

V. DEMONSTRATIONS

The following trials were carried out in the open-source driving simulation environment OpenDS[3], which is available online in the Dreams4Cars Zenodo community [28]. To run the tests on a desktop PC, and make them more accessible, the pedals and the steering wheel have been emulated with keyboard arrows (gaming control devices are also supported).

The examples demonstrate emergent interactions, showing that 1) the human driver can influence the co-driver agent's

[3]https://opends.dfki.de/

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/LRA.2021.3136646, IEEE Robotics and Automation Letters

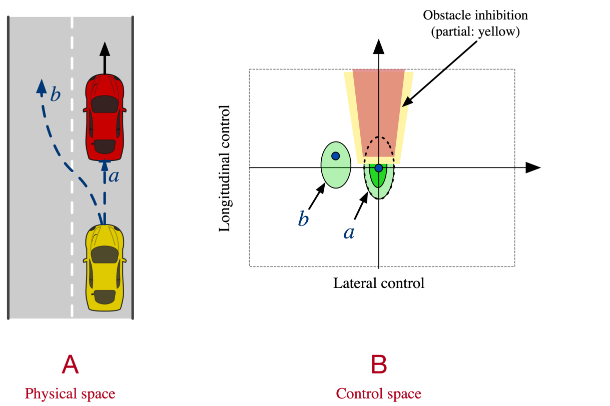

Fig. 4. (A) Bird's-eye view of an example scenario with two affordances: *a* "car-following" and *b* "overtaking". (B) In the control space, active regions (green) and inhibited regions (yellow or red) encode the value of lateral and longitudinal control [3, Section II.C.5].

decision-making, obtaining a lane change whenever possible and safe; 2) the co-driver can dismiss the human's suggestions if they are useless or dangerous.

A. Example 1: Overtaking by acting on the steering wheel

This example shows how a human can promote an overtake maneuver by signaling the intention with the steering wheel. The scenario is described by Fig.4, A: a two-lane straight road where overtaking is possible and safe. The leading vehicle (red car) travels at $45\ \mathrm{km/h}$, and the speed limit is $50\ \mathrm{km/h}$. In this condition, the co-driver agent opts for car following because the added value of overtaking would not compensate for the cost of maneuvering to another lane. The situation is described in Fig.4, B, where the value of action a slightly prevails over action b even if the latter would be a little faster (see [3, Section III-C] for an in-depth discussion on this point).

Note that even if the co-driver agent chooses action a it is aware that another affordable action b exists. The latter is instantiated in the motor space: it remains covert but ready to be used. A little lateral bias as the one presented in equations (5)-(6), may hence be sufficient to switch action selection from safe action a to safe action b. In this process, human activity (slightly steering leftwards) is interpreted and projected onto the leftwards actions that the co-driver agent recognizes in the environment.

Figure 5 shows the biases produced by the human driver and the resulting interaction: the lateral bias, in the "Lateral" sub chart, is the normalized steering angle α of equation (6); the longitudinal bias is $(g - b)$ of (4).

When the human driver repeatedly presses the left arrow on the keyboard (corresponding to a leftward rotation of the steering wheel), at circa 310 m, the lateral driver bias (α) gradually increases, as shown by the red dotted line. At approximately 350 m, the bias becomes strong enough to switch the action selection to b, which is shown in the chart with the black line indicating that the co-driver intention becomes a left lane change. Following the beginning of the lane change, at circa 360 m, the human stops tapping the left arrow (equivalent to keeping the steering wheel in the rotated position). At circa 390 m, the lane change is completed (green dashed line). Then, even if the human keeps the leftward bias, the co-driver agent switches to keeping the (current) lane because there is no other lane to the left. At about 420 m, the human releases the left arrow (equivalent to releasing the steering wheel). Later, at about 590 m, after passing the slow car, the co-driver autonomously returns to the right lane (because in the autonomous mode, w_R is slightly greater than w_L to prefer the right lane if free). The bottom chart of Fig.5 shows the longitudinal state of the co-driver: "car-follow" means that the choice of the longitudinal control is limited by inhibitions of the preceding vehicle, such as the situation a in Fig.4 B, whereas "free-flow" means free choice of longitudinal control, e.g., the situation b.

For the sake of clarity, Fig.6 shows the trajectory and the steering wheel angle δ (magnified by a factor of 10).

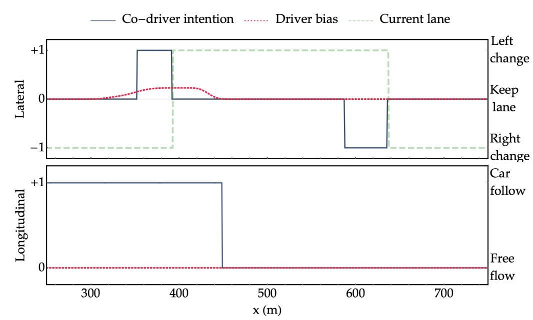

Fig. 5. Example 1. Human bias (red) and co-driver agent actions (black). The leftward bias maps onto an existing alternative action (Fig.4 *b*) that is adopted by the co-driver agent.

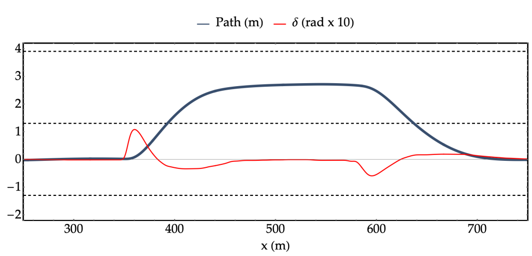

Fig. 6. Example 1. Trajectory of the co-driver car and steering wheel angle.

Note that the co-driver agent reacts dynamically: it continuously updates action priming and action selection (in our implementation, every 50 ms). Thus, if human drivers change their minds, they can specify a new biasing pattern steering the co-driver decision-making within one loop (50 ms). For example, suppose they suddenly feel uncomfortable and want to stop. In that case, they may tap the back arrow (or touch the brake pedal), negatively increasing the longitudinal bias, effectively signaling to choose actions that reduce the speed (stop). Aborting the overtaking maneuver could also be obtained by tapping the right arrow at any time.

B. Example 2: Overtaking by pressing the gas pedal

The next example complements the above, considering the same scenario but using the longitudinal bias. In this case, the bias (g) is produced by pressing the gas pedal (i.e., repeatedly pressing the up arrow in the OpenDS version), which gradually increases the weight of the bias function given in (4).

Figure 7 presents the results. After the human presses the gas pedal at around 390 m, the longitudinal bias increases, as shown by the red dotted line in the bottom chart of Fig.7. At approximately 420 m, the longitudinal bias is strong enough to switch the co-driver selection to b, as shown by the co-driver intention on the top chart. The lane change is completed shortly after. The maneuver lasts a shorter time than in Section V-A because the human driver —with the pressure of the gas pedal— explicitly requests acceleration; whereas, in the previous example, the human driver was asking for a lane change (which permits a faster speed after the obstacle is cleared). After the overtaking maneuver, the co-driver agent returns autonomously to the right lane, which is logical given that the right lane is free. However, as in Example 1, the human driver can oppose the return to the right with actions that keep the salience of the left lane high.

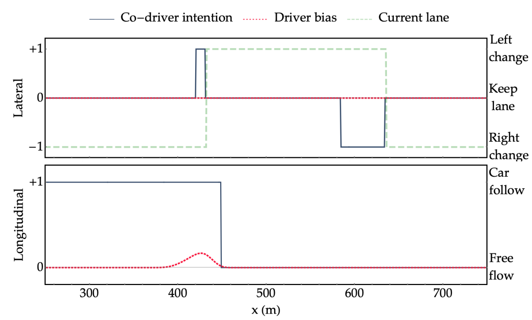

Fig. 7. Example 2. Human bias (red) and co-driver agent actions (black). The forward bias maps onto an existing alternative action (Fig.4 b) that is adopted by the co-driver agent.

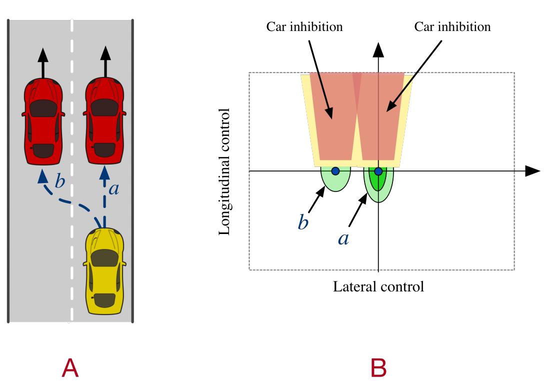

Fig. 8. Alternative scenario. The co-driver (yellow car) has two possible goals: either follow the car on the left a or that on the right b. There is no long-term velocity convenience in changing lane.

C. Example 3: Steering wheel pulse with blocked left lane

In this case, the left lane is occupied by a second car, as shown in Fig.8, A. Both cars travel at 45 km/h. If not blocked, the co-driver could travel at the the speed limit, i.e., at 50 km/h. There are two possible behaviors that are instantiated in the motor space with two salience functions, humps and local maxima (blue dots): a and b. Unlike Fig.4, b is not faster than a —the two have the same longitudinal control. Furthermore the salience of a is greater than b because the latter discounts the cost of the lane change. Without human directives the co-driver hence follows car a.

Figure 9 reports what happens when the human driver steers to the left, at around 400 m. Although choosing the left lane does not carry any tangible speed improvement, it can be safely carried out, and the co-driver correctly understands the human's directive and moves to the left lane. As shown in the bottom chart, the agent remains in the car-following state.

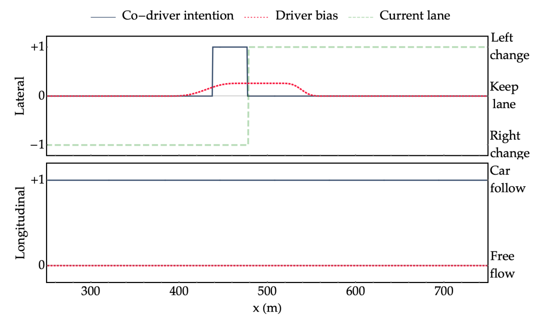

Fig. 9. Example 3. Human bias (red) and co-driver agent actions (black). The leftward bias maps onto an existing alternative action (Fig.8 b) producing a lane change without overtaking.

D. Example 4: Gas pedal pulse with blocked lanes (rejected)

This case uses the same scenario above (Fig.8) but, instead of turning the steering wheel, the human driver presses the gas pedal.

In this example, whatever the weight, the longitudinal bias acts uniformly on the local maxima b and a. The co-driver agent does not change its intention because the human faster speed request cannot be satisfied. Fig.10 shows that even after pressing the gas pedal to full stroke (the red dotted line raises to 1 between 400 m and 580 m in the bottom chart), there is no effect on the co-driver agent behavior. However, if the vehicle on the left lane were slightly faster, there would be an effect.

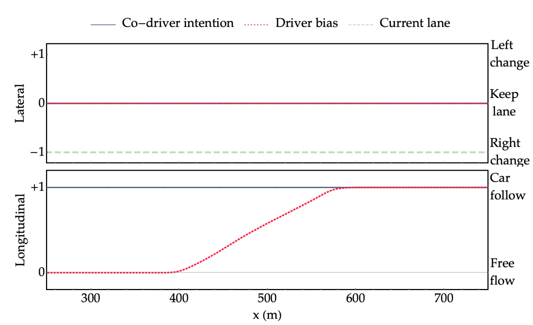

Fig. 10. Example 4. Human bias (red) and co-driver agent actions (black). The forward bias does not maps onto any existing alternative action (Fig.8) and is ineffective.

E. Example 5: Vetoed overtake request

The last example presents the principle of the lower-level veto [3, Section II.E-1]. Figure 11 A shows the scenario: the

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/LRA.2021.3136646, IEEE Robotics and Automation Letters

co-driver agent follows car a while the left lane is entirely blocked by an overtaking car. In the control space (B), the local maximum corresponding to goal b no longer exists: the salience hump for b is completely zeroed by the inhibition of the second car.

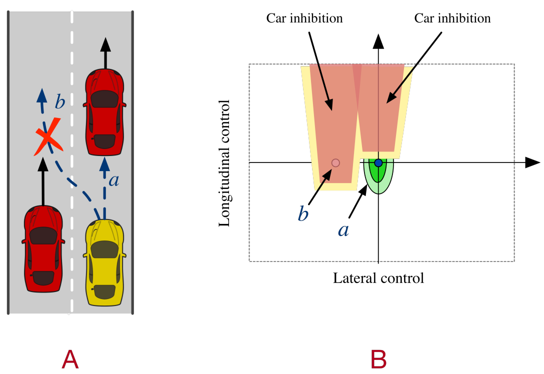

Fig. 11. Alternative Example 2. The co-driver can only follow car a. Lane change b is blocked by an overtaking car.

In the example, the human driver attempts to force a lane change by completely turning the steering wheel to the left, as shown in the top chart of Fig.12. This input raises the weight of the left bias (Fig.4) to 100 % (red dotted line), effectively doubling the salience of any local maxima in the left lane. However, this has no effect because no affordable action exists in that lane. Hence, the co-driver vetoes the human's indication and remains on the right in the car-following condition. If the human had pressed the gas pedal, the same result would have been obtained.

Note that if at any moment the car on the left shows courtesy and slows down, then the inhibition in the control space will be updated. As soon as b is cleared, the co-driver will accept the lane change request and comply with the human (i.e., no lane change occurs until the car on the left gives way).

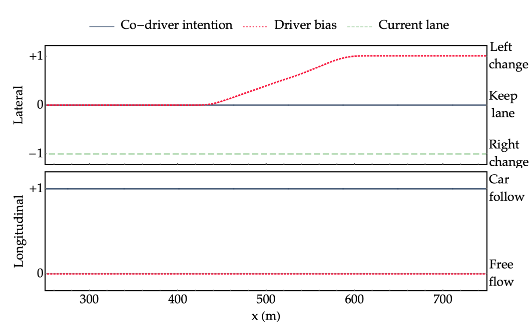

Fig. 12. Example 5. Human bias (red) and co-driver agent actions (black). The leftward bias does not maps onto any existing alternative action (Fig.8) and is rejected

VI. GENERALIZATIONS AND LIMITATIONS

The previous examples dealt with the same two-lanes road. However, the agent's sensorimotor organization (Section III) can accommodate a variable number of affordances, as they are primed in parallel. For example, if there were a bifurcation (i.e., two diverging roads), there would one affordance of type a in Fig.1 per road. Similarly, if there were other lanes, there would be further affordances of type c susceptible to biases, with corresponding, non-coincident activations in the motor space. For example, if there were a further lane c at the left of b in Fig.8 A, there would be a further activated region c in Fig.8 B. A longitudinal bias could hence cause the selection of c and a double lane change would occur. If the lane marking between b and c were solid, c would not be a legal affordance and would not be primed into the motor space, preventing the crossing of the solid line. On the other hand, if double lane change maneuvers were considered dangerous in the highway code and prohibited, action priming could be restricted to the adjacent lanes. This does not mean that the agent cannot drive on multilane roads but that it can only change one lane at a time.

With the inhibition mechanism of Fig.2 the direction of the traffic may be arbitrary: first, future space-time regions are determined from path prediction, and then they are used for inhibition of the corresponding motor space regions regardless of their order. This also works in intersections.

In the supplementary material, we provide three videos that demonstrate more general situations such as the above. Video 1 shows an intersection scenario with two-way traffic. It compares the behavior of the unbiased co-driver and the co-driver under the human request of traveling faster. It shows that the co-driver accepts to cross, using a partially inhibited affordance with fewer safety margins. Video 2 shows how the human driver can tell the codriver to stop in a free parking spot using a combination of backward and rightward biases. Finally, Video 3 shows driving in a winding two-way road under the human request of traveling as fast as possible on the left lane (maximum forward and leftward bias). The co-driver returns to the right lane when meeting vehicles in the opposite direction. At some point, the co-driver aborts an overtake maneuver initiated under the strong human bias.

A. Limitations of this work

The goal of this paper is to demonstrate mirroring via biasing. The study implicitly assumes that the interaction is between two peer agents, mainly that the affordances instantiated by the codriver are safe (hence the codriver can reject unsafe requests). This may be true for SAE automation levels 4 and 3 as long as the operative condition meets the operational design domain. However, the human driver might wish (or need) to take some tighter control in various circumstances. One primary reason for allowing a tighter control modality stems from the ethical principle that states that artificial intelligence should never hinder the human will. Another practical reason for allowing tighter control may be related to the fact that even a fully automated level 4 system might miss unconventional affordances that could be useful in some circumstances. For example, if the road were blocked, a level 4 car might stop forever. However, a human driver could be aware that gravel on the side of asphalt is "drivable" (an affordance that a level 4 vehicle might miss) and might carefully drive offroad to pass the blockage.

The switch to tight control could be triggered when the human biases exceed given thresholds. The implementation of unrecognized affordances can be obtained with a *modified action-priming* mechanism: the agent assumes that the human might see affordances it does not know, and the whole motor space is instantiated with a uniform value so that the agent will copy the human actions. The agent can keep partial inhibitions for the obstacles to resist actions leading to a certain collision.

A different case is when an imperfect co-driver is unsure about the safety of some of the affordable actions. For example, a level 2 system might be confident about the safety of in-lane maneuvers but not about lane change. The implementation of interactions when the agent is unsure of specific affordances may be obtained with a *modified action selection* mechanics: the uncertain affordances (e.g., the lane change) may be chosen only if they are simultaneously biased so that an implicit confirmation is obtained. We must remark that the above considerations are research directions for future work (not studied in this paper).

VII. CONCLUSIONS

We have shown how emergent smart interactions between humans and intelligent vehicles of the type postulated by the "H-metaphor" can be obtained by biasing the decision-making of the co-driver agent. For this, the co-driver agent must have a sensorimotor architecture modeled on the affordance competition hypothesis. Human activity is used to set weights that influence the co-driver action selection to navigate the affordance landscape. The human driver input is a dummy driving control that exploits the traditional human control devices available on passenger cars, i.e., the steering wheel and pedals. The work presented here is a study that nicely shows the working principle, including the rejection of unsafe commands.

Future work is planned in a driving simulator to collect the human's opinion regarding the interaction with the agent (how the agent can be influenced and how humans perceive the agent response). Additional work will concern the definition of learning-based biasing functions (e.g., (4)-(6)).

REFERENCES

[1] F. Flemisch, C. Adams, S. Conway, K. Goodrich, M. Palmer, and P. Schutte, "The H-Metaphor as a Guideline for Vehicle Automation and Interaction," *Control*, no. 2, pp. 1–30, Dec. 2003.

[2] D. Norman, *The Design of Future Things*. 387 Park Avenue South, New York, NY, USA: Basic Books, 2007.

[3] M. Da Lio, R. Donà, G. P. R. Papini, and K. Gurney, "Agent architecture for adaptive behaviors in autonomous driving," *IEEE Access*, vol. 8, pp. 154 906–154 923, 2020.

[4] M. Da Lio, R. Donà, G. Rosati Papini, F. Biral, and H. Svensson, "A mental simulation approach for learning neural-network predictive control (in self-driving cars)," *IEEE Access*, vol. 8, pp. 192 041–192 064, 2020.

[5] G. P. R. Papini, A. Plebe, M. D. Lio, and R. Donà, "A reinforcement learning approach for enacting cautious behaviours in autonomous driving system: Safe speed choice in the interaction with distracted pedestrians," *IEEE Transactions on Intelligent Transportation Systems*, pp. 1–18, 2021.

[6] K. H. Goodrich, P. C. Schutte, F. O. Flemisch, and R. A. Williams, "Application of the h-mode, a design and interaction concept for highly automated vehicles, to aircraft," in *2006 ieee/aiaa 25TH Digital Avionics Systems Conference*, 2006, pp. 1–13.

[7] B. Vanholme, D. Gruyer, S. Glaser, and S. Mammar, "A legal safety concept for highly automated driving on highways," in *2011 IEEE Intelligent Vehicles Symposium (IV)*, 2011, pp. 563–570.

[8] F. Flemisch, Y. Canpolat, E. Altendorf, G. Weßel, M. Itoh, F. Flemisch, M. Baltzer, M.-P. Pacaux-Lemoine, D. Abbink, and P. Schutte, "Shared and cooperative control of ground and air vehicles: Introduction and general overview," in *2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, 2017, pp. 858–863.

[9] M. Da Lio, F. Biral, E. Bertolazzi, M. Galvani, P. Bosetti, D. Windridge, A. Saroldi, and F. Tango, "Artificial Co-Drivers as a Universal Enabling Technology for Future Intelligent Vehicles and Transportation Systems," *IEEE Transactions on Intelligent Transportation Systems*, vol. 16, no. 1, pp. 244–263, 2015.

[10] G. Csibra and G. Gergely, "'Obsessed with goals': functions and mechanisms of teleological interpretation of actions in humans." *Acta psychologica*, vol. 124, no. 1, pp. 60–78, 01 2007.

[11] A. N. Meltzoff, "The 'like me' framework for recognizing and becoming an intentional agent." *Acta Psychologica*, vol. 124, no. 1, pp. 26–43, 2007.

[12] J. Gibson, "The theory of affordances," in *The Ecological Approach To Visual Perception*. Lawrence Erlbaum Associates, 1986, pp. 127–143.

[13] S. Hurley, "The shared circuits model (SCM): how control, mirroring, and simulation can enable imitation, deliberation, and mindreading." *The Behavioral and brain sciences*, vol. 31, no. 1, pp. 1–58, 02 2008.

[14] G. Hesslow, "The current status of the simulation theory of cognition." *Brain research*, vol. 1428, pp. 71–9, Jan. 2012.

[15] M. Jeannerod, "Neural simulation of action: a unifying mechanism for motor cognition." *NeuroImage*, vol. 14, no. 1 Pt 2, pp. S103–9, 07 2001.

[16] M. Haruno, D. M. Wolpert, and M. Kawato, "Mosaic model for sensorimotor learning and control." *Neural Computation*, vol. 13, no. 10, pp. 2201–2220, 2001.

[17] D. M. Wolpert, K. Doya, and M. Kawato, "A unifying computational framework for motor control and social interaction." *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, vol. 358, no. 1431, pp. 593–602, 03 2003.

[18] Y. Demiris and B. Khadhouri, "Hierarchical attentive multiple models for execution and recognition of actions," *Robotics and Autonomous Systems*, vol. 54, no. 5, pp. 361–369, 05 2006.

[19] Y. Demiris, "Prediction of intent in robotics and multi-agent systems." *Cognitive processing*, vol. 8, no. 3, pp. 151–8, 09 2007.

[20] P. Cisek and J. F. Kalaska, "Neural mechanisms for interacting with a world full of action choices," *Annual review of neuroscience*, vol. 33, pp. 269–298, 2010.

[21] P. Redgrave, T. J. Prescott, and K. Gurney, "The basal ganglia: a vertebrate solution to the selection problem?" *Neuroscience*, vol. 89, no. 4, pp. 1009–1023, 1999.

[22] R. Bogacz and K. Gurney, "The basal ganglia and cortex implement optimal decision making between alternative actions," *Neural computation*, vol. 19, no. 2, pp. 442–477, 2007.

[23] L. F. Nunes and K. Gurney, "Multi-alternative decision-making with non-stationary inputs," *Open Science*, vol. 3, no. 8, p. 160376, 08 2016.

[24] G. Pezzulo and P. Cisek, "Navigating the Affordance Landscape: Feedback Control as a Process Model of Behavior and Cognition," *Trends in Cognitive Sciences*, vol. 20, no. 6, pp. 414–424, 06 2016.

[25] T. J. Prescott, P. Redgrave, and K. Gurney, "Layered Control Architectures in Robots and Vertebrates," *Adaptive Behavior*, vol. 7, no. 1, pp. 99–127, 01 1999.

[26] P. Bosetti, M. Da Lio, and A. Saroldi, "On the human control of vehicles: an experimental study of acceleration," *European Transport Research Review*, vol. 6, no. 2, pp. 157–170, 2014.

[27] A. Plebe, M. Da Lio, and D. Bortoluzzi, "On Reliable Neural Network Sensorimotor Control in Autonomous Vehicles," *IEEE Transactions on Intelligent Transportation Systems*, vol. 21, no. 2, pp. 711–722, Feb. 2020.

[28] M. Da Lio, "Dreams4Cars Horizon 2020 project community on zenodo," 10 2020. [Online]. Available: https://zenodo.org/communities/dreams4cars/