MSPRT action selection model for bio-inspired autonomous driving

Gastone Pietro Rosati Papini
Dept. of Industrial Eng.
University of Trento, Italy
gastone.rosatipapini@unitn.it

Giammarco Valenti
Dept. of Industrial Eng.
University of Trento, Italy
giammarco.valenti@unitn.it

Alice Plebe
Dept. of Industrial Eng.
University of Trento, Italy
alice.plebe@unitn.it

Mauro Da Lio
Dept. of Industrial Eng.
University of Trento, Italy
mauro.dalio@unitn.it

***Abstract*—This paper proposes a bio-inspired action selection mechanism, the multi-hypothesis sequential probability ratio test (MSPRT), as a decision making tool in the field of autonomous driving. We investigate the capability of the MSPRT algorithm to effectively select the optimal action whenever the autonomous agent is required to drive the vehicle. We present numerical simulations to demonstrate the robustness of the MSPRT action selection when dealing with noisy measurements.**

I. INTRODUCTION

Autonomous vehicles (AVs) require effective algorithms to perform robust decision making in the shortest time frame possible. In a dynamic environment such as the one faced by the AVs, the capability of reacting promptly is a major factor in potentially avoiding collisions and saving lives. The inherent complexity of the process is worsened by the presence of sensors' noise and uncertainties, which affect the way the behavioural level selects the proper action.

Several theories have been proposed in the literature on how animals perform effective decision making [1]. For instance, in [2] the *affordance competition* concept underlines a parallel processing of multiple actions competing against each other until the selection of the winning behavior. Such a modeling framework is based on the definition of criteria for assessing the *worthiness* of the action and the *selection* process itself.

We exploit this concept of parallel competing actions in the context of the European Projects SafeStrip[1] and Dreams4Cars[2]. In particular, in SafeStrip we take advantage of the mirroring mechanism introduced in [3], [4] to infer the human driver intended action. Such an inference process boils down to the selection among a set of optimality-based longitudinal maneuvers, called *motor primitives*, of the one matching the driver intended action in terms of instantaneous jerk j_0. In Dreams4Cars we utilize a similar optimality-based motor primitives approach for the synthesis of an autonomous driving agent called Co-driver [5]. In addition to the longitudinal manoeuvres, we also generate set of lateral manoeuvres by defining a 1-dimensional grid on instantaneous lateral jerk r_0. By combining the two grids we devise a 2-dimensional matrix called *motor cortex* where each entry is a pair of (j_0, r_0) which encodes a latent action. Each pair is then assigned a merit via the definition of a scenario dependent *salience*.

[1]https://www.safestrip.eu
[2]https://www.dreams4cars.eu

The rest of this paper is devoted to demonstrate how we can perform such a task taking advantage of a biologically inspired action selection mechanism.

II. THE MOTOR CORTEX CONCEPT

To better clarify how the affordances competition process takes place, let us inspect an example simulation scenario as in Fig. 1. In the proposed situation the ego car, driven by the Co-driver agent, is travelling at high speed on a straight road when a slower vehicle is detected.

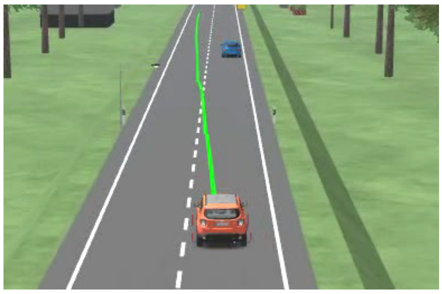

Fig. 1: Example of simulation scenario in bird-eye view

The *motor cortex* (at time t indicated with $\mathcal{M}_t$) encodes the *affordances* and it can be computed by introducing some merit criterion. For the considered example scenario we model the merit as the maximum time at which, given the pair (j_0, r_0), the vehicle will leave the road or collide with other road users (*minimum intervention principle* [6]). By establishing the criterion above, we can compute an artificial motor cortex as in Fig. 2, where the salience is displayed along the z-axis of the 3D plot. It can be noticed how lateral controls close to zero have high merit values, while steering abruptly has a close to zero salience. Each of the action compete against the others for winning the selection process. The outcome of the "competition" is the optimal pair (j_0^*, r_0^*) that will eventually guide the car for the next time-step.

III. ACTION SELECTION

A. WTA algorithm

The most trivial approach to model the affordances competition would be to simply choose the pair having the highest instantaneous salience. This selection mechanism is known as

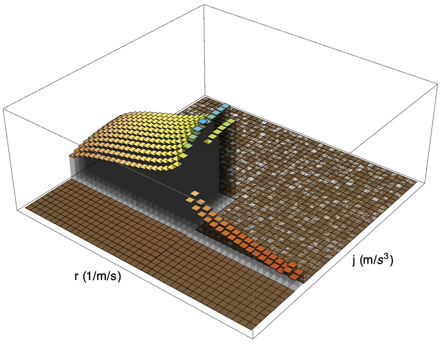

Fig. 2: Motor cortex based on the minimum intervention principle.

winner takes all (WTA) [1] and has proven to be fairly efficient in the simulation environment where there is no signal noise.

B. MSPRT algorithm

In order to overcome the WTA limits in case of noise, we propose here the multi-hypotheses sequential probability ratio test (MSPRT) [7] decision making algorithm. The key idea of the MSPRT algorithm is to accumulate *evidence* for each channel and then pick an action only when the integral reaches a predefined *threshold* level. The MSPRT has been shown to be asymptotically time-optimal in a multi alternatives process [8] just like the basal ganglia of the human brain [9].

The procedure for action-selection using the MSPRT algorithm is the following:

Result: Action log-likelihood
$\mathcal{M}_{\text{list}} \leftarrow \mathcal{M}_t$;
$\bar{\mathcal{M}}(t) \leftarrow \text{mean}\,\{\mathcal{M}_{\text{list}}\}$;
$L(t) = \bar{\mathcal{M}}(t) - \log \sum_{i,j} \exp\left(\bar{\mathcal{M}}_{ij}(t)\right)$;
if $\max(\exp(L)) >$ *threshold* **then**
  take action;
  $\mathcal{M}_{\text{list}} = \lambda\ \bar{\mathcal{M}}(t)$
else
  follow previous action;
end

If some of the channels computed by $\max(\exp(L))$ reaches a predefined *threshold* value, we take the action with maximum evidence, otherwise we continue to follow the previous action. The overall behaviour of the MSPRT algorithm can be shaped by adjusting the hyper-parameters: *threshold* = 0.0005 (slows down the switch to a new channel), *windows size* = 8 (average out noise), λ = 0.9 (forgetting factor, introduces a memory effect).

IV. Simulation Comparison

We compare the performance of the MSPRT against the WTA on simulated logged data. Firstly we let the agent drive on a simulated scenario with no noise affecting the measurements. According to this set-up we can perform optimal decision making using a simple WTA algorithm. We then select a 9-seconds long critical double lane change maneuver where the responsiveness of the action selection plays a fundamental role. Next, we re-execute the simulation offline, *i.e.* we take the logged motor cortex history, we apply some random noise on the channels and we re-execute the decision making algorithm only on the corrupted motor cortex. We then analyze again the performances of the WTA against MSPRT with respect to the ground-truth case obtained previously.

Fig. 3 reports the results of assessment as a function of the adimensional noise variance σ injected into the motor cortex. In case of limited noise figures, the WTA still outperforms MSPRT due to the worse transient performance of the latter. As soon as we introduce noise in the simulation, however, the advantages of the MSPRT start to be evident.

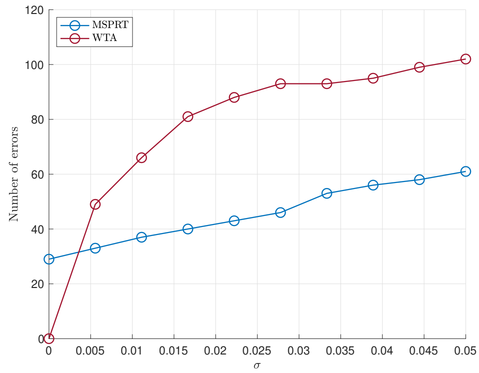

Fig. 3: MSPRT vs. WTA channels selection errors.

Acknowledgements

This work is supported by the European Commission Grant 731593 (Dreams4Cars) and 723211 (SafeStrip).

References

[1] 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.

[2] P. Cisek, "Cortical mechanisms of action selection: the affordance competition hypothesis," *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 362, no. 1485, pp. 1585–1599, 2007.

[3] M. Da Lio, A. Mazzalai, and M. Darin, "Cooperative Intersection Support System Based on Mirroring Mechanisms Enacted by Bio-Inspired Layered Control Architecture," *IEEE Transactions on Intelligent Transportation Systems*, vol. 19, no. 5, pp. 1415–1429, May 2018. [Online]. Available: https://ieeexplore.ieee.org/document/8011474/

[4] G. Valenti, D. Piscini, and F. Biral, "A cooperative intersection support application enabled by safe strip technology both for c-its equipped, non-equipped and autonomous vehicles," in *European conference on Intelligent Transportation Systems*, 2019.

[5] 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.

[6] E. Todorov and M. I. Jordan, "A minimal intervention principle for coordinated movement," in *Advances in neural information processing systems*, 2003, pp. 27–34.

[7] C. W. Baum and V. V. Veeravalli, "A sequential procedure for multihypothesis testing," *IEEE Transactions on Information Theory*, vol. 40, no. 6, 1994.

[8] V. Draglia, A. G. Tartakovsky, and V. V. Veeravalli, "Multihypothesis sequential probability ratio tests. i. asymptotic optimality," *IEEE Transactions on Information Theory*, vol. 45, no. 7, pp. 2448–2461, 1999.

[9] 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.