Dream-like simulation abilities for automated cars

DREAMS4CARS

Grant Agreement No. 731593

Deliverable: D3.2. – Simulation system (release 2)

Dissemination level: CO – Confidential

Delivery date: 29/October/2019

Status: Final

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 731593

Deliverable Title	Simulation system (release 2)		
WP number and title	WP3 Dream-like states		
Lead Editor	Henrik Svensson, HIS		
Contributors	Mauro Da Lio with help from team members: Riccardo Donà, Alice Plebe, Gastone Pietro Rosati Papini, Francesco Biral, Daniele Bortoluzzi, Paolo Bosetti, Luca Zaccarian, Enrico Ber-tolazzi UNITN		
	Alex Blenkinsop, Sean Anderson, USFD		
	Seyed Ali Ghorashi, David Windridge, MU		
	Henrik Svensson, Sara Mahmoud, Erik Lagerstedt HIS		
	Rafael Math, DFKI		
Creation Date	17/April/2019	**Version number**	1.3
Deliverable Due Date	31/October/2019	**Actual Delivery Date**	29/October/2019
Nature of deliverable			
	X	DEM – Demonstrator (software prototype with report)	
Dissemination Level/ Audi-ence			
	X	CO - Confidential, restricted under conditions set out in MGA	

Version	Date	Modified by	Comments
0.1	14/April/2019	Mauro Da Lio	Initial draft (initial document structure and contributions to sections 1 and 3).
0.2	29-31/May/2019	Mauro Da Lio	Enhanced document organization. Contribu-tions to sections 1.1, 2.1, 2.2.
0.3	5/July/2019	Henrik Svensson	Small contributions to 2.3 and added de-tailed organization of section 4. Updated the list of contributors.
0.4	26/August/2019	Mauro Da Lio	Almost completed sections 3.1 and 3.2.
0.5	30/August/2019	Mauro Da Lio	Initial writing of section 3.3
0.6	20/September/2019	Mauro Da Lio	Section 5.3 (introduction part)

0.7	25/September/2019	Mauro Da Lio	Section 3.1.9. Improved section 3.2. Sections 3.3.1, 3.3.2, 3.3.3.
0.8	01/October/2019	Gastone Pietro Rosati Papini	Integration of contribution concerning Reinforcement Learning. Sections 4.1, 4.3
0.9	01/October/2019	Alice Plebe	Section 5.1.1
0.9.2	09/October/2019	Henrik Svensson	Integration of pedestrian models (New section 2.4 and 2.4.1 mostly by Rafael Math and 2.4.2 with contributions from Erik Lagerstedt) Minor corrections to chapter 3. Additions to Section 4 and 5. Integration of contribution concerning progressive learning (New Section 4.2.2), pedestrian models (New Section 4.2.3; by Rafael Math, 4.3.3.2: contributions by Erik Lagerstedt), and experiments with occlusions (New Section 4.2.5, contributions from Sara Mahmoud)., more pedestrian types (New Section 4.2.6, contributions from Sara Mahmoud), and higher-level RL (New Section 4.2.7 mostly by Alex Blenkinsopp).
0.9.3	18/October/2019	Mauro Da Lio	Sections 3.3.4 to 3.3.7
1.0	21/October/2019	Henrik Svensson	Minor corrections throughout the document. Moved pedestrian models to Chapter 4. Additions to Chapter 2 and 4. Rewrites and additions to section 4.2.6 and 4.
1.1	25/October/2019	Henrik Svensson	Minor changes due to internal review comments by Sean Andersson and Serge Thill. Integration of section 5.1.2 (written by David Windridge and Seyed Ali Ghorashi). Additions to 1.1.
1.2	28/October/2019	Mauro Da Lio	Final version for submission
1.3	29/October/2019	Henrik Svensson	Quality improvements to the text in 4.4.2, 4.4.4 and 4.4.5.

Executive Summary

This deliverable represents the main research findings of Dreams4Cars, i.e., a process for bootstrapping sensorimotor systems based on *learning models of the wold* (both models that predict the effect of actions and models that predict events) *that are then manipulated* offline to *synthetize* action strategies (Figure 1).

This process is described in section 3. It starts with the supervised learning of vehicle dynamics "forward" (or predictive) models from data. We here introduce an efficient architecture for the networks that makes the models explainable (thus solving the black box issue that would otherwise prevent adoption of neural network for motion control due to liability issues and the ISO 26262 mandatory requirement). We show how this architecture allows disentangling independent causes (weak superposition effect) and how it is very similar to cerebellar filters. Stochastic forward models are then introduced resorting to the bootstrapping technique which does not make any assumption about the distributions and correlations of data. Progressive refinement of models via lifelong learning (e.g., learning nonlinearities on top of linear models) is also explained. A particular form of overfitting (high-frequency spectral overfitting) is studied and methods of regularization are provided. Statistical methods to compare predictions of learned forward models are finally introduced.

In section 3.2 the learning of inverse models is introduced showing two possible approaches: from data (supervised learning by swapping input and output) and via the first form of episodic simulations (unsupervised learning). We show that the latter is generally more robust, and it may be used to train inverse model for robust predictive control (an aspect that is still a research topic for traditional MPC).

In section 3.3 we step to a higher level of motor control introducing the notion "short-cuts" in the simulation paths. This is a form of learning neural network abstractions that allows making prediction of action outcomes without needing to simulate the entire action in details. It can be regarded as the transition between embodied (detailed) and episodic (abstract) simulations. This in turns allows to progressively build more and more abstract, fast and efficient simulation blocks for simulations of further higher levels, ending with the tools for accelerating Reinforcement Learning.

We finally end up with the learning of action values, i.e., the salience stored in the "motor cortex" and we show how it is functionally equivalent to the notion of reward in Reinforcement Learning, except that it is obtained via a synthesis process that manipulates learned models of the world rather than via trial and error exploration (which is the RL way).

In section 4, Reinforcement learning applications are introduced. The choice of safe speed (as the most important behavioural choice for safe driving) is studied in details for the case of pedestrians possibly crossing the road; obtaining a network that interacts with the lower-levels of the agent via setting recommended safe speed. Open issues related to RL are also discussed (RL in itself is not the research focus of Dreams4Cars).

We introduce also possible research lines for the future in section 5; in particular how the sensorimotor system here developed can be integrated with self-organizing perception systems to form a whole. We take inspiration from the CDZ hypothesis of Damasio and show some proof of concept with implementations based on Variational Autoencoders.

Across all the document we make as many efforts as possible to compare the methods here developed with more or less traditional alternatives, and point many advantages: so, section 3.1.5 evaluates the performance of forward models networks, section 3.1.9 deals with statistical comparisons of forward models, section 3.3.4 compare stochastic motor models in the literature, section 3.3.6 compares to Optimal Control, section 3.3.7.2 makes an internal comparison between the core bootstrapping approach of section 3 and Reinforcement Learning. Section 4 also presents some qualitative comparisons of the RL implementation within similar ones in the literature. Worth to be noted is the fact that here RL creates networks that are built on top of an agent that can operate on the real world, whereas the literature examples are for simulated worlds and are not directly transferable to the real world (transfer here is permitted by acting with high-level directives on an agent that can already operate on the real world by itself)