25th ITS World Congress, Copenhagen, Denmark, 17-21 September 2018

Paper ID EU-TP1320

A System for Human-like Driving Learning

Mauro Da Lio [1*], Alice Plebe [2*], Daniele Bortoluzzi[1],
Gastone Pietro Rosati Papini[1], Riccardo Donà[1]

1. Department of Industrial Engineering, University of Trento, Trento, Italy
2. Department of Information Engineering and Computer Science, University of Trento, Trento, Italy
* Email: {mauro.dalio, alice.plebe}@unitn.it

Abstract:
One of the major advances expected for future societies is the development of autonomous driving systems able to assist humans in one of the most common daily tasks of people all over the world. An impressive interest, involving both producers and research institutes, has grown in the last decade. Within the broad variety of directions taken by researchers towards autonomous driving systems it is possible to identify a common underlying paradigm, that of the sense-think-act scheme. Even if it has been the ground for an important progress in several components of a driving agent, it is exposed to several serious problems. One of the most dramatic issue is the demand of testing in order to qualify a level of reliability of an agent. This quality assessment requires a huge experimental campaign. In order to alleviate this problem, we propose a different approach which better adheres to the human driving strategy, as revealed by contemporary cognitive sciences. In this paper we describe the overall architecture that is under development in the framework of the EU Horizon 2020 Dreams4Cars Research and Innovation Action, in which the artificial agent inherits the broad structure of the human motor control strategy, as known by the state-of-the-art in cognitive science.

Keywords:
Automated Driving, Cognitive Systems.

Introduction

Recent years have seen an increasing enthusiasm in the amount of industrial and research activities aimed at implementing automated driving. However, the main challenge that has yet emerged is how to demonstrate that engineered vehicles are safer than humans.

A recent study [1] has shown that to provide *statistically significant* evidence that an autonomous driving agent is safer than human drivers, it would be necessary to test drive for billions of miles, which is not feasible. To date, companies testing autonomous driving vehicles in the State of California have accumulated less than 0.5 million miles in approximately one year [2]. The problem arises from the fact that humans are very reliable at driving: in the United States there are only about 1,09 deaths and 77 injuries per 100,000,000 miles.

This position paper proposes a paradigm shift for the architecture of driving agents. Today, most of autonomous driving (AD) approaches follow the *sense-think-act* paradigm, which means that the perception system produces a symbolic representation of the environment and that there is software that determines the agent's behavior based on this representation: this is known as Cartesian rationalism. This approach almost necessarily means that the software must be fully developed and tested by human designers. Simulations can be used to reduce the amount of testing but the tested scenarios, once again, must be known. For this kind of architecture, it is definitely a challenge to obtain a system provably capable of error-free behavior for billions of miles.

Such a system should be able to act in *autonomy*, i.e. in situations that might not have been expressly considered at the design stage. Humans have the ability to rely on the few generic situations faced within a few hours of driving schools, to deal with novel situations for the rest of their driving life. This is possible also because they continuously learn new behavioral strategies from all the potentially critical situations they can experience, in real life or even only imagined.

It should be noted that human accidents are largely due to distractions, fatigue, risk-taking or driving under the effects of alcohol or drugs. Such aspects obviously do not affect artificial systems. We, therefore, want to draw inspiration from mechanisms that underlie human autonomy and propose an artificial cognitive architecture with similar autonomy, able to increase its own skills as senior drivers do. For this reason, we analyze in the rest of the paper the human sensorimotor system and show how it could be taken as model for the architecture of driving agent.

How the Brain Learns to Drive a Car

As reviewed in the Introduction, human beings are surprisingly brilliant and reliable at driving, even if this ability is certainly outside of our natural endowment. What we are endowed with, instead, is a extraordinary capability to learn highly specialized and complex sensorimotor behaviors in general. The same flexible control system in our brain that during early development learns how to walk, run, and jump, can later learn complex motor coordination such as playing piano, skying, playing tennis, and, driving cars and motorcycles [3].

No artificial system has so far come close to replicating the superlative learnability of the sensorimotor system of the human brain. It is in the tradition of Artificial Intelligence the attempt to exploit the principles with which the human brain accomplishes difficult tasks in designing similar artificial systems. However, most current research lines for automatic driving systems seems to have ignored this path, and this is a project that attempts a design that broadly follows the way in which humans arrive at the ability to guide.

The task of driving is solved by the human brain with the same kind of strategy we adopt for every type of motor planning that requires a continuous and complex perceptual feedback. Surely the operations of this general strategy are far from being fully understood and there are several different competing theories, without unanimous consensus. However, progress achieved in the understanding of the neurocognitive processes involved in one of these performance is remarkable, and such as to make it possible to grasp some useful general guidelines for the design of artificial car control systems.

The neurocognitive foundation of driving adopted here falls within the simulative theory of cognition, of which an extensive description is found in [4]. In short, the idea is that cognitive activity is performed, at least in some cases, as a simulated interaction with the environment. In more details, simulation for Hesslow is a general principle of cognition explicated in three different ways: in the simulation of behaviour, the simulation of perception, and anticipation. Cues of all of these three forms of simulations give grounds for our design approach [5]. The first form, simulation of behaviour, means that the signal coming from the prefrontal cortex through pre-motor areas may activate the motor cortex even if it is stopped before producing overt behaviour. A simulated behaviour is therefore essentially a deleted or incomplete action. The advantages of this strategy for the brain is the possibility to mentally engage with concepts related with behaviours.

A second form, simulation of perception, correspond to sensorial imagery, a widely studied phenomenon, especially in the visual modality. Sensor imagery is the mechanism by which a representation of the type created during the initial phases of perception is present, but the stimulus is not actually being perceived [6]. Mental imagery has a strong connection with the phenomenon of dreaming, that also activates the lower visual areas with neural patterns not much different from online visual processing. There are several evidences that lead to claim that, during the first childhood, dreaming is fundamental to build up simulative skills [7]. It is not yet fully clear the mechanism by which the low visual areas can be activated either by signals from the thalamus or through upper cortical areas. One of the best explanations points to cortical components called *convergence-divergence zones* (CDZs) [8]. These components receive convergent projections from the early sensorimotor areas and send back divergent projections to the same areas. This arrangement has the first purpose to record the combinatorial organization of the knowledge fragments coded in the lower visual areas, together with the coding of how those fragments must be combined to represent an object as a whole. CDZ records are gradually learned by experience with objects in the environment. The CDZ framework can explain perceptual imagery, as it proposes that similar neural networks are activated when objects or events are processed online and when they are recalled from memory.

How Simulation is Realized

Today, the simulative theory of cognition finds its strength in a series of empirical evidences of its actual implementation in the brain. An overview of suche evidences, together with a comprehensive framework on how simulation might be realized in the brain, is found in Cisek [9, 10]. Simulation appears as the main component of the mechanisms by which the brain selects actions and specifies the parameters or metrics of those actions. Cisek's theory is named *affordance competition hypothesis*, in which "affordance" is the term originally proposed by Gibson [11] to refer to the action possibilities of the environment that are available to an organism. For example, for most western humans a knife "affords" cutting food, and a fork "affords" forking it; and for a driver a lane "affords" either lane change or lane following and a yield line "affords" either stopping or crossing. In Cisek, "affordances" are the internal representations of the potential actions which are in constant competition for deciding the next behavior. In this project, the term affordances will be used in exactly the same sense.

The first brain part involved in the affordance competition hypothesis is the occipito-parietal *dorsal stream*. In the traditional division of the visual processing path into the *dorsal stream* and the occipitotemporal *ventral stream* the former builds a representation of *where* things are and the latter of *what* things are [12]. An updated account of the dorsal stream proposes that its role is to mediate various visually guided actions by several substreams. For example, the lateralintraparietal (LIP) area specify potential saccade targets; the medial intraparietal (MIP) area specify possible directions for arm reaching.

The next fundamental bain part is made up by the basal ganglia in connection with the dorsolateral prefrontal cortex (DLPFC), and performs action selection. Since action selection is a fundamental problem faced by even the most primitive of vertebrates, it is consistent with the involvement of an ancient structure conserved throughout evolution, like basal ganglia. DLPFC appears to play the role of collecting "votes" for categorically selecting one action over others. There is a wide literature on the detailed mechanisms by which basal ganglia and prefrontal cortex interact in taking decisions [13–16].

The last fundamental component is the cerebellum, where an internal predictive feedback is generated, once the final selected action is released. The importance of this brain region is highlighted by the fact that it contains the majority of brain neurons in all mammals, and up to 80% of all neurons in the human brain. One of the earlier proposals including the cerebellum in simulation is the theory of neural emulators [17], that bridges a close link between the engineering domain of control theory and signal processing to neural representations. In terms of control theory, Grush's emulators are essentially forward models, and are located in the cerebellum. There are several neuroscientific evidences for forward models in the cerebellum [18, 19], but the issue is controversial. According to the analysis done in [20] several assumptions about the cerebellum in terms of engineering schemes for motor control and signal processing appear to be biologically implausible. This include the Kalman filter assumed by Grush, the Smith predictor, and the feedback-error-learning scheme for adaptive inverse control. Also schemes like forward models, according to this analysis, appear unlikely to map directly into cerebellar microcircuits. What is suggested is that, instead, cerebellar microcircuits learn a task-specific adaptive-filter operation. However, from recent physiological and morphological evidences it seems that at least the cerebro-cerebellum, the phylogenetically newer part of the cerebellum, provides a forward model for limb movement [21].

Overall Architecture

The system that will mimic aspects of human drive learning has an overall architecture described in this section. Three different modes of working are planned for this system: the simulative mode, the real driving mode, and the quality assurance mode. Fig 1 shows an outline of the system.

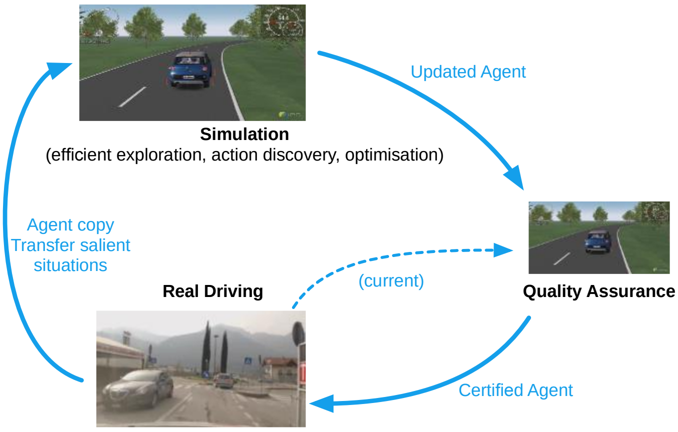

Figure 1 - The system architecture, consisting of three components: real driving, simulation/optimization and quality assurance.

The Simulative Mode

This is the mode where knowledge of human drive learning is mostly leveraged. The agent in this mode explores a simulated world and learns new behaviors, by testing different situations. Several alternatives are planned for the implementation of this mode. A first one is based on the open-source virtual reality driving simulation environment OpenDS (www.opends.eu), that implements multibody systems (MBS) technology [22], with some machine learning extensions. MBS is a general way for engineers to simulate large-scale physical system dynamics, including complex systems of bodies under the action of external forces, control loops, and other conditions. One significant advantage of MBS, compared to alternatives, is longer-term accuracy deriving from the universal physical principles substrate.

MBS can accept a large number of parameters representing physical quantities. Selecting the model parameters from a distribution can be a method for generating imaginary conditions. Nevertheless, the distribution values must be plausible (it does not make sense to simulate a road with a friction coefficient equal to 10 since this never happens in reality). Thus, the distributions can either derive from the agent's observations whilst in the real driving mode or from a-priori knowledge, when it is is possible without loss of generality.

Genetic algorithms can be also used to create recombinations of elements in a simulation, e.g. swapping a car with a pedestrian or a stop behaviour with a cross behaviour. This way a genetic algorithm can mimic recombination of percepts that occur in dreams. Moreover the algorithm can be biased for searching situations that are useful for learning.

Another alternative approach for generating imaginary situations is to use the same deep neural network model at the core of the inverse model of the agent, that will be described in sec. 4.1. This model has one principal decoding part that generates detailed motor commands from features at low dimensions, but also a secondary decoding part decoding back to perceptual signals. The combination of the coding part with this secondary decoder works as autoencoder, therefore it can generate a number of imaginary situations by randomly exploring its feature space, an exploration that can be constrained towards those situations that are more likely to be useful during real driving.

A further possibility is to adopt generalized, top-down “motor babbling”, a well established notion in the framework of perceptionaction machine learning [23, 24] and applied in cognitive robotics [25]. This method derives its name from the analogy with the process of motor learning in infant humans: starting from minimal sets of percepts and actions, the entire action space is randomly sampled. For each new motor action that produces a discernible perceptual output in the current perceptual set, the produced percept is allocated. This process can be carried out at various levels in the perception-action hierarchy.

The final purpose of the simulative cognition at all levels is the development of optimized behaviors. There are analytical ways to compute solutions within the *optimal control* (OC) domain, a methodology for producing optimal solutions for goal-directed problems [26]. Despite being OC a methodological framework based on differential algebraic equations and other mathematical structures, there are empirical evidences that the planning of hand movement to reach a target, as performed by the brain, corresponds to the minimization of the integral of the squared jerk, as in OC [27]. Therefore, OC too is in line with the brain-inspired design principle of this project.

Another possible approach is *motivated learning*, inspired by the learning mechanisms in animals, and thus better suited for higher-level strategies. In this case, the agent takes advantage of opportunistic interactions with the environment to develop a knowledge of what actions cause predictable effects in the environment. In this way, the agent is able to build internal models of action-outcome pairings in the brain.

The Real Driving Mode

This mode corresponds with real driving, during which the agent records situations that are considered worthy to be re-enacted in the simulative mode. Those salient situations are determined according to several criteria. The most relevant events occur when there is a discrepancy between the predictions of the internal model and what actually happens in the real world. These situations can point out some imperfections in the prediction/planning model, or are the indication of the occurrence of a novel condition the agent has no knowledge about. Other relevant criteria that are taken into account are space-time separation with other vehicles, jerkiness of control, compliance with traffic rules, and also traffic and energy efficiency. In order to be able to re-run experienced situations, the agent notes any event considered worthy of further analysis. The agent logs both the low-level sensory and control signals, and the high-level signals represented by the internal states of the agent's architecture. In such way, it is possible to record the "intentions" of the agent (or the estimated intentions of the human driver) at all levels of the sensorimotor system. Hence, the agent is able to run simulations of alternative lower-level strategies while preserving the higher-level intentions.

Quality Assurance Mode

Lastly, the third mode shown in Fig. 1 has the purpose of quality assurance. As the simulative mode, this mode is based on a multibody simulation system too, this one called CarMaker, developed by IPG Automotive (ipg-automotive.com).

By comparing several test cases, the system in this mode certifies new versions of the agent, ensuring that any updated agent does not function worse than the previous optimum. It also helps to identify situations of over-fitting, when an agent learns to better cope with the most recently-dreamed situation only, at the expenses of earlier ones.

Architecture of the Driving Learner

The overall architecture of the agent here proposed is largely derived from the neural components and brain strategies that allow humans to drive, as described in the previous sections. For this reason, the main scheme of the agent, shown in Fig. 2, is superimposed on a sketched brain, and is an adaptation to the driving task by an artificial agent of the action selection scheme in [10, Fig. 1 p. 278].

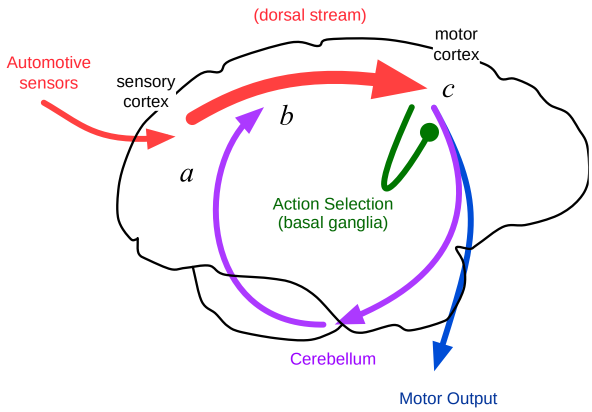

Figure 2 - Main scheme of the agent architecture, the main pathways are depicted in red (dorsal stream), green (action selection) and violet (forward model); (a), (b), and (c) are progressive stages of representations in a scale from sensorial to affordances.

In the remaining of this section the three main processing pathways shown Fig. 2 will be described.

Inverse Model

In the outlook given by Cisek the inverse model is one of the role by the cortical visual dorsal flow, by transforming sensorial signals into motor signals that may cause the perceived stimuli. This is the sort of transformation that in our agent correspond to the red pathway represented in the Fig. 2.
Coherently, the input of the transformation are sensory data in (a), that take typically the format of topographic representation of perceptual space, merging information from sensors such as the camera and LIDAR. The format in (b) will be in a topology that is more oriented towards an egocentric control space, in two dimensions, corresponding to the degrees of freedom in driving: lateral and longitudinal controls. The final transformation occurs in (c), in a format that, following Cisek, we call *affordances*, in the sense of meaningful trajectories in the two dimensional motor control space, that can represent actions like changing a lane, stopping at a red traffic light, and so on. The pathway from (a) to (c) is not necessary encapsulated, there is the possibility of intermediate insertion of inputs from intelligent sensors capable of providing data with format like that in (b). In addition, it will not be implemented as a single transformation, rather as a collection of specialized substreams, as in fact is the case in the cortical visual dorsal pathway. Possible substreams encompass:

- road geometry;
- obstacles;
- traffic lights or other traffic directives.

Note that while road geometry will provide the active regions in the space of longitudinal and lateral controls, obstacles and traffic lights will produce as output inhibited regions.
Several of the substreams from (a) to (c) will be based on neural network models, broadly inspired by the CDZ [8] described in sec. 2.1. The main similarity that these neural models share with the CDZ of Meyer and Damasio is the compression/decompression of dimensionality. A first network reduce from high dimension sensory inputs down to the final low dimensional space. This final space, in turn, is used as input to a neural network that acts as a decoding in which the dimensions of the levels are gradually increasing. When the desired output of the decoding network is the reconstruction of the original input, the chain of the two networks implements the *autoencoder*, the idea of Hinton that triggers the invention of deep learning [28, 29]. The self-encoder has thus become one of the most effective methods for generating new data with deep learning [30]. We will use our networks for both purposes: as decoder for generating motor controls, and as autoencoder for generating reconstructed perceptual data playing the role of imagery.

Action Selection

This component approximately behaves as the basal ganglia in the brain, typically in coordination with DLPFC, as described in sec. 2.2. It takes decision on the agent’s behavior, and can take into account high-level biases, such as the level of automation or driving style. Its implementation will be based on multihypothesis sequential probability ratio test (MSPRT), an asymptotically optimal statistical test for decision making, which has been shown to be a possible computational abstraction of decision function performed by the cortex and the basal ganglia.

Forward Model

This component corresponds roughly to the processes supposed to be performed in the cerebellum, as discussed in sec. 2.2: forward model. Its input is composed of efferent copies of motor commands, and the output is in terms of intermediate perceptual data of the scene, as if the motor commands were issued. This is how it would be possible to identify salient and potentially dangerous situations without actually experiencing them in real driving.

Conclusions

This paper presents a vehicle control architecture which novelty with respect to other approaches is the grounding in the human driving mechanisms. This objective is accomplished by structuring the agent in the way that best resemble the structures of the brain when involved in driving sensorimotor control, for what is currently known in cognitive science. A distinctive feature of this agent is the mechanism for learning perceptions-action through experience, even in dream-like simulations. We believe that this approach may offer advantages, both in terms of resources needed for development, and for quality assessment.

Acknowledgment

The EU Horizon 2020 Dreams4Cars Research and Innovation Action project is supported by the European Commission under Grant 731593.

References

1. Kalra, N., & Paddock, S. M. (2016). Driving to safety: How many miles of driving would it take to demonstrate autonomous vehicle reliability?. *Transportation Research Part A: Policy and Practice*, *94*, 182-193.
2. Dixit, V. V., Chand, S., & Nair, D. J. (2016). Autonomous vehicles: disengagements, accidents and reaction times. *PLoS one*, *11*(12), e0168054.
3. Wolpert, D. M., Diedrichsen, J., & Flanagan, J. R. (2011). Principles of sensorimotor learning. *Nature Reviews Neuroscience*, *12*(12), 739.
4. Hesslow, G. (2012). The current status of the simulation theory of cognition. *Brain research*, *1428*, 71-79.
5. Da Lio, M., Mazzalai, A., Thill, S., Svensson, H., Gurney, K., Anderson, S. R., ... & Heich, H. J. (2017). Exploiting Dream-Like Simulation Mechanisms to Develop Safer Agents for Automated Driving. In *ITSC 2017*. IEEE.
6. Moulton, S. T., & Kosslyn, S. M. (2009). Imagining predictions: mental imagery as mental emulation. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1521), 1273-1280.

7. Thill, S., & Svensson, H. (2011, January). The inception of simulation: a hypothesis for the role of dreams in young children. In *Proceedings of the Annual Meeting of the Cognitive Science Society* (Vol. 33, No. 33).
8. Meyer, K., & Damasio, A. (2009). Convergence and divergence in a neural architecture for recognition and memory. *Trends in neurosciences*, *32*(7), 376-382.
9. Cisek, P. (2007). Cortical mechanisms of action selection: the affordance competition hypothesis. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *362*(1485), 1585-1599.
10. Cisek, P., & Kalaska, J. F. (2010). Neural mechanisms for interacting with a world full of action choices. *Annual review of neuroscience*, *33*, 269-298.
11. Gibson, J. J. (1979). *The ecological approach to visual perception: classic edition*. Psychology Press.
12. Ungerleider, L. (1982). Two cortical visual systems. *Analysis of visual behavior*.
13. Redgrave, P., Prescott, T. J., & Gurney, K. (1999). The basal ganglia: a vertebrate solution to the selection problem?. *Neuroscience*, *89*(4), 1009-1023.
14. Doya, K. (1999). What are the computations of the cerebellum, the basal ganglia and the cerebral cortex?. *Neural networks*, *12*(7-8), 961-974.
15. Bogacz, R., & Gurney, K. (2007). The basal ganglia and cortex implement optimal decision making between alternative actions. *Neural computation*, *19*(2), 442-477.
16. Lewis, J. M., Chambers, J. M., Redgrave, P., & Gurney, K. (2011). A computational model of interconnected basal ganglia-thalamocortical loops for goal directed action sequences. *BMC neuroscience*, *12*(1), P136.
17. Grush, R. (2004). The emulation theory of representation: Motor control, imagery, and perception. *Behavioral and brain sciences*, *27*(3), 377-396.
18. Wolpert, D. M., Miall, R. C., & Kawato, M. (1998). Internal models in the cerebellum. *Trends in cognitive sciences*, *2*(9), 338-347.
19. Jeannerod, M., & Frak, V. (1999). Mental imaging of motor activity in humans. *Current opinion in neurobiology*, *9*(6), 735-739.
20. Porrill, J., Dean, P., & Anderson, S. R. (2013). Adaptive filters and internal models: multilevel description of cerebellar function. *Neural networks*, *47*, 134-149.
21. Ishikawa, T., Tomatsu, S., Izawa, J., & Kakei, S. (2016). The cerebro-cerebellum: Could it be loci of forward models?. *Neuroscience research*, *104*, 72-79.
22. Blundell, M., & Harty, D. (2004). *The multibody systems approach to vehicle dynamics*. Elsevier.
23. Shevchenko, M., Windridge, D., & Kittler, J. (2009). A linear-complexity reparameterisation strategy for the hierarchical bootstrapping of capabilities within perception–action architectures. *Image and Vision Computing*, *27*(11), 1702-1714.
24. Windridge, D. (2017). emergent intentionality in Perception-Action subsumption Hierarchies. *Frontiers in Robotics and AI*, *4*, 38.
25. Dearden, A., Demiris, Y., Kaelbling, L. P., & Saffotti, A. (2005). Learning forward models for robots. *International Joint Conference on Artificial Intelligence*.
26. Bertolazzi, E., Biral, F., & Da Lio, M. (2005). Symbolic–numeric indirect method for solving optimal control problems for large multibody systems. *Multibody System Dynamics*, *13*(2), 233-252.
27. Liu, D., & Todorov, E. (2007). Evidence for the flexible sensorimotor strategies predicted by optimal feedback control. *Journal of Neuroscience*, *27*(35), 9354-9368.
28. Hinton, G. E., & Salakhutdinov, R. R. (2006). Reducing the dimensionality of data with neural networks. *science*, *313*(5786), 504-507.
29. Hinton, G. E., Osindero, S., & Teh, Y. W. (2006). A fast learning algorithm for deep belief nets. *Neural computation*, *18*(7), 1527-1554.
30. Dosovitskiy, A., Springenberg, J. T., Tatarchenko, M., & Brox, T. (2017). Learning to generate chairs, tables and cars with convolutional networks. *IEEE transactions on pattern analysis and machine intelligence*, *39*(4), 692-705.