Reinforcement Learning for Urban Autonomous Driving

**Student:** Carlos Fernando Coelho Valadares (cfcv@cin.ufpe.br)

**Advisor:** Hansenclever de França Bassani (hfb@cin.ufpe.br)

**Domain:** Autonomous driving

08 September 2020.
1. Abstract

While humans seem to handle well the urban driving task, it is still a big challenge for autonomous vehicles. Most of the current autonomous driving software deployed in real cars consists of a modular pipeline. This pipeline provides useful information about how the driving decisions are being made but makes it hard to replicate the intended vehicle behavior as each submodule has to be optimized separately. Recently, with advances in deep learning, end-to-end optimization approaches started to raise. It consists in using a single deep learning model trained via Imitation Learning or Reinforcement learning to replace the entire modular pipeline, making it easier to optimize and deploy with the cost of decreasing the system interpretability. In this work, we explore the potential of end-to-end optimizations with reinforcement learning to tackle the autonomous driving task. The soft actor-critic algorithm is applied with different sets of observations in order to discover which features are important to succeed in this task. Furthermore, we propose a novel framework for tackling the autonomous driving task in urban scenarios introducing the notion of interpretability on end-to-end reinforcement learning, which consists in connecting path planning algorithms with reinforcement learning controllers. We will evaluate this new framework against the best approaches on the CARLA Leaderboard Challenge on simple navigation scenarios.

Enquanto humanos aparentam não ter muitos problemas para aprender a dirigir, essa tarefa ainda é um grande desafio para os carros autônomos. O software utilizado, atualmente, por carros autônomos é baseado em um pipeline modular. Este pipeline fornece informações importantes sobre como as decisões estão sendo tomadas, porém, ele torna difícil de replicar o comportamento desejado ao veículo, pois cada submodule deve ser otimizado separadamente. Recentemente, com os avanços em Deep Learning, otimizações end-to-end começaram a surgir. Esse método consiste em substituir o pipeline modular por um modelo Deep Learning, deixando mais fácil a otimização, mas com um custo de diminuir a interpretabilidade do sistema. Neste trabalho, exploraremos o potencial de otimizações end-to-end juntas com aprendizagem por reforço para o problema da direção autônoma. O algoritmo Soft Actor-Critic é aplicado com diferentes conjuntos de observações para descobrir quais features são importantes para ter sucesso nesta tarefa. Além disso, propomos uma nova estrutura para abordar o problema da direção autônoma em ambientes urbanos, o qual introduz a noção de interpretabilidade em otimizações de aprendizagem por reforço end-to-end. Esta nova abordagem consiste em conectar algoritmos de planejamento de caminhos com controladores do tipo aprendizagem por reforço. Este novo método será avaliado e comparado com as abordagens que obtém as melhores performances no CARLA Leaderboard Challenge em cenários de navegação simples.
2. Introduction

In the last decades, a lot of work has been done to integrate new technologies into vehicles in order to increase driving comfort as cruise control and park assistance. Efforts have been made to produce vehicles as autonomous as possible, hoping to provide more safety on the streets and accessibility for people that cannot drive. However, driving in urban areas is a really hard task and although humans seem to handle it well, the driving task is still a big challenge for the automotive industry. The main approach currently used to drive vehicles autonomously is called modular pipeline. It is composed of submodules where each module has a well-defined task as: perception, prediction, path planning, and control. As this approach brakes down high-level tasks into a pipeline, it provides good interpretability of the system and information about how the driving decisions are being made. On the other hand, it is very hard to tune each module separately and then put everything together to replicate the intended vehicle behavior.

In order to overcome this problem, end-to-end approaches were conceived to replace the entire modular pipeline. It receives raw sensor data as input and outputs directly the control values of the vehicle. These types of approaches are not new, works in the field of end-to-end autonomous driving can be seen in the 80’s or 90’s as [1], [2], however limited by the technologies back then, these type of approaches did not move forward. With recent advances of deep learning, these end-to-end approaches started to rise again, at this time using deep learning models trained via Imitation Learning [3]–[6] or Reinforcement Learning [7]–[10] to replace the modular pipeline. These approaches provide an easier way to optimize the pipeline, as the sub-modules will not need to be tuned separately. Nevertheless, compared to the modular pipeline approach, it is much less interpretable as the model acts as a black box.

In this work, we propose end-to-end methods using reinforcement learning for tackling different autonomous driving scenarios as car racing and urban driving. Our studies focus on discovering which features are important for the agent to succeed in these tasks and the introduction of the notion of interpretability in end-to-end methods.
3. Objectives

The general objective of this work is to study which basic features are important to the learning algorithms drive in a simple scenario and to introduce the notion of interpretability in end-to-end reinforcement learning methods for urban driving scenarios.

The specific objectives of this work are:

• Find which features are important for driving using end-to-end reinforcement learning pipelines;
• Create a latent space for accelerate training;
• Apply a proposed implementation to a simple simulated urban driving scenario.
• Propose a new framework for end-to-end reinforcement learning to the urban driving scenario with the notion of interpretability;
• Compare the results with state of the art methods.
4. Methodology

The first step for the development of this work is the literature review. As this work encompasses two fields, a state of the art study will be needed for both reinforcement learning and autonomous driving domains. This will allow us to choose a state of the art reinforcement learning algorithms to tackle the autonomous driving task.

Once the literature is revised, the chosen reinforcement learning algorithm will be applied to the driving task in an end-to-end way. A simple car racing scenario will be simulated for the experiments. This part will focus on finding which features are important to the reinforcement learning agent for learning to drive properly a vehicle. For that reason, the agent will be trained with several modifications in the observation and action spaces.

Finally, the CARLA simulator [11] will be used to simulate the urban driving scenario. With the knowledge gathered from the literature and the experiments in the car racing scenario, an end-to-end reinforcement learning approach will be proposed for tackling the urban driving task. This approach will introduce the notion of interpretability in end-to-end reinforcement learning methods and its performance will be compared to state of the art approaches in simple navigation scenarios.
5. Schedule

<table>
<thead>
<tr>
<th>Activity</th>
<th>Sept</th>
<th>Oct</th>
<th>Nov</th>
</tr>
</thead>
<tbody>
<tr>
<td>Literature review</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Algorithm study</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Experiments</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Analyse of results</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dissertation writing</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Presentation preparation</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
</tbody>
</table>
References


6. Possible Evaluators

The possible evaluators are the professors:

- Germano Crispim Vasconcelos (gev@cin.ufpe.br)
- Cleber Zanchettin (cz@cin.ufpe.br)
- Tsang Ing Ren (tir@cin.ufpe.br)
7. Signatures

Hansenclever de França Bassani
Advisor

Carlos Fernando Coelho Valadares
Student