**Type of project (e.g. Bachelor thesis, Master AI thesis, Master OR thesis, Master AI internship project, Master OR internship project, Master AI semester project, Master OR semester project):**

Master AI semester project

**Year of project:**

2018-2019

**Key words (5 max):**

reinforcement learning, deep learning, latent space, transfer learning, neural networks.

**Name(s) of student(s) who participated in the project:**

Adrian Rodriguez Grillo, Danni Liu, Alessandro Scoppio, Kevin Trebing, Tonio Weidler.

**Name(s) of supervisor(s):**

Dr. Kurt Driessens.

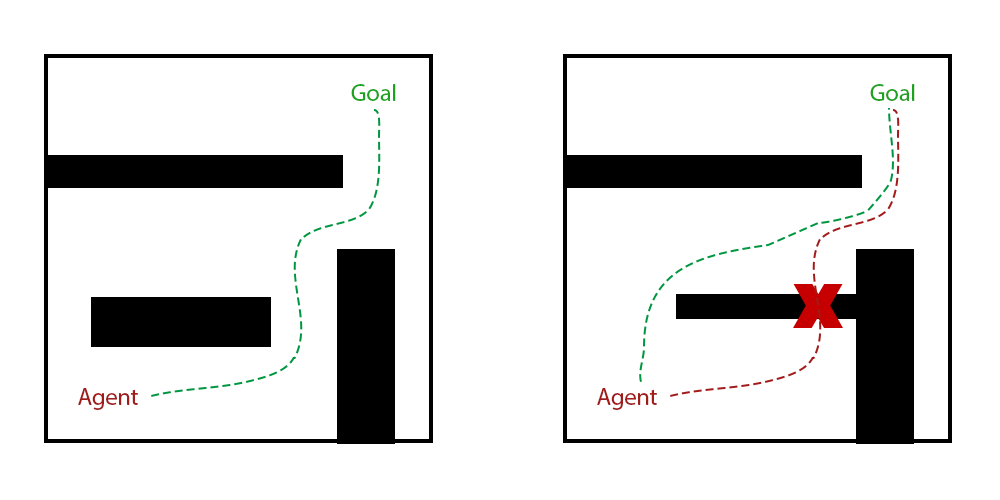
**Title of the project:**

Reinforcement Learning in Latent Space.

**Download link for additional content (at least 1 image/figure, report, final presentation, maybe some movies):**

<https://drive.google.com/open?id=1KERu1x7GRxv9AhhZ5xR4yotCyZN2pkNy>

**Captions of images/figures. Enter your images (in low resolution if needed) here as well with appropriate captions:**

Figure 1: An agent learning a path from its current location to a goal on the left map. When applying the learned knowledge to the map on the right, the path is insufficient. Instead, the agent should learn an abstract concept of avoiding obstacles and reaching a goal.

**Problem statement and motivation. Please provide a general description of the problem that you worked on, a short description of why this is relevant, and a few sentences that state precisely what you aimed at:**

**Reinforcement learning** is a subfield in artificial intelligence that shares many commonalities with human learning. In reinforcement learning, a virtual agent is set free in an unknown environment to solve a task, for example walking towards a destination. Whilst exploring, it will receive rewards or punishments depending on its behavior. These are used to reinforce desired or suppress undesired behavior. In the example of reaching a destination, this could be whether it moves towards or away from it. The agent will act via trial-and-error, attempting to collect the maximum possible reward. By repeatedly doing different kinds of actions in the environment, the agent *explores* and thereby accumulates knowledge about what to do given a particular situation. The agent measures the quality of an action based on its expectations about the rewards it would lead to in the future. The behavior that the agent learns is called a *policy,* which describes at every step what he should do. After its learning phase, the agent should have inferred the best behavior from the rewards and punishments it received.

Unlike humans and animals, the virtual agent can explore the unknown environment very quickly, thanks to fast computer simulations. However, when facing a new task, it cannot take advantage of previously learned knowledge and must explore and learn again from scratch. In this project, the goal is to enable the agent to use previously learned knowledge on a new task, or **transfer learning.** Transferring knowledge is what allows humans and animals to solve problems they never faced before and quickly adapt to them. Imagine a human that learns how to build a chair from scratch. If he masters this skill, we can expect him to quickly find out how to build a table, applying the knowledge he gained from building chairs. A virtual agent cannot do this without expensive and task-specific help from its creator. In fact, an essential part of the learning capabilities of humans is being able to construct abstractions, or knowledge representations. Neural networks are mathematical models designed to construct such abstractions. The abstraction is often referred to as **latent space**, because it is supposed to reveal characteristics of the input that are not directly visible. This latent representation would eventually embed general knowledge about the environment. When the agent learns how to behave in an environment based on abstracted information, we hope that it will be able to utilize the gained knowledge in new tasks by applying the same abstraction technique.

Imagine a virtual agent that needs to find a way from its current location to a destination, but this time, there are obstacles blocking the direct path. The agent can easily solve this problem by learning for each position on the map what action brings it closest to the target. However, only remembering the best path will not help it, when we change the obstacles on the map. If a new obstacle blocks the original solution, the agent will never reach its goal. See figure 1 for an illustrative description of this situation. What we want to achieve, is that the agent learns to abstract its visual input. Instead of remembering one path, it should learn the concept of tracing around objects and moving in the direction of the goal. Optimally, this would make the arrangement of obstacles irrelevant and the agent becomes capable of finding its way - even on unknown maps.

We aim at finding a way of guiding the abstraction process such that it both contains information specific enough to solve learned tasks but also general enough to transfer the knowledge to new tasks. We will investigate and compare multiple approaches to this problem. To evaluate them, we collect sets of tasks that are similar enough to allow knowledge transfer. We conduct first experiments on very easy problems but extend our scope to more difficult and potentially diverse tasks as soon as we achieve good results on the former.

Being able to reuse previously learned knowledge would help to advance the field of reinforcement learning by potentially making training an agent more efficient, due to its capability of leveraging knowledge from previous tasks. This could make it possible to solve tasks that are for now very ineffective to solve.

**Research questions and/or hypothesis: explain the questions/hypothesis that you addressed during your project:**

In the following research questions, *learning* refers to effectively exploring an unknown environment and accumulating knowledge about which actions are good to take in given situations.

1. How can a virtual agent find an abstraction method that captures common characteristics of several similar yet still different tasks, and then learn to perform all these tasks only based on the single abstract description?
2. After learning a task using the abstract description, can the virtual agent more easily learn to perform a previously unseen but similar task (e.g. using less time)?
3. After learning a task using the abstract description, can the virtual agent directly perform a previously unseen but similar task without learning?

**Major outcomes of your project. Please provide up to seven major outcomes of your project. These should be formulated in a clear but compact way:**

* We will propose an approach to allow an agent finding an abstract way to represent different task environments, which still allows the agent to learn.
* We expect to find that knowledge about different tasks can be partially preserved. More specifically, this should be achievable in two steps:
  + Find an abstract description that applies to these tasks.
  + Translate task-specific knowledge via the common abstract description.
* We intend to deliver:
  + A software capable of training agents to perform a selected range of tasks.
  + A virtual agent that is trained on a set of tasks but can complete another task without being trained on it.
  + Answers to our research questions supported by experiment results and analysis.

**References: Provide possible references you are using in your text in the APA standard. Use e.g. Google Scholar to obtain references in the correct format:**

**Any additional information you would like to be mentioned:**

**To be filled by supervisors (dear supervisors: delete what you do not want):**

◻ I agree that this project is posted on the DKE student projects webpage. (Feel free to make any modifications to the text/material provided by your students.)

◻ I agree that the final project report is uploaded to the DKE student projects webpage.

◻ I agree that the final presentation is uploaded to the DKE student projects webpage.

◻ I agree that movies provided by the students are uploaded to the DKE Youtube channel and shown on the DKE student projects webpage.