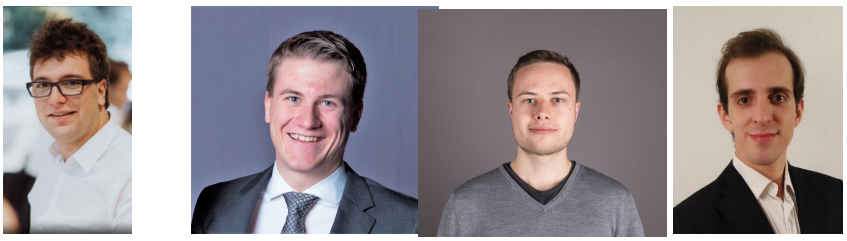
# Navigation By Reinforcement Learning

3D Vision Project Proposal
Supervised by: Nicolay Savinov
March 9, 2018


GROUP MEMBERS

Varin Buff | Gian Heinrich | Björn Joos | Robin Stähli




## I. DESCRIPTION OF THE PROJECT

In the Industry 4.0 the need for automation increases tremendously and thus the demand for robots with high navigation abilities in complex three-dimensional surroundings grows vastly. Especially deep learning allows an efficient approach to achieve this. Savinov *et al.* [3] used an animal inspired non-metric [2] [1] semi-parametric topological memory (SPTM) approach. With this approach, they were able to improve their success rate by a factor of three compared to their baselines. Our team will evaluate further tensorforce based baselines which follows the same evaluation procedure as introduced by [3] and will allow to put their work on an even more solid foundation.

## II. WORK PACKAGES AND TIMELINE

Timeline and project planning Our project is organized in four main parts. We will start with an orientation phase to gain some knowledge about previously done work, focused mainly on Nikolay Savinovs' paper [3]. We will then use our gathered insights to deploy the same environment GITHUB for training and testing of our agent within the vizdoom environment GITHUB and use the same evaluation methods of Reinforcement Learning baselines that were used in [3]. We go for this setup to be later able to benchmark the performance of our agent against the results of the paper. Our vizdoom agent will be trained with the AC3 algorithm REFCITE and if there is time left also with PPO REFCITE. To guarantee an efficient workflow we divide our team into two subgroups, where one team is entrusted with the infrastructural needs of the group, e.g. deploy necessary software on Leonhard. The rest of the group will be focused on the Reinforcement Learning algorithm training and evaluation.
We will devide our team in two subgroups to tackle the following tasks:

- Orientation: Understand previously done work by Nikolay Savinov our Tutor and deploy a similar setup, e.g. use the same Vizdoom version and adapt the training and evaluation methods of RL baselines used in his paper.
- Interface Vizdoom envioment
- Study Agent movement
- Deploy test program to Leonhard
- Understand and implement A3C (and eventually PPO)
  - Implement rewards and training structure to explore efficiently the maze
  - Training: we train our agent with the A3C algorithm
  - Validation: We use the validation mazes to tune our parameters
  - Testing: testing will be done with the seven provided mazes
- If there is time left we train and test our agent with PPO in addition
- We compare our results to previous results of Niklay Savinov
- Write the final report

## III. Outcomes and Demonstration

The expected outcome of this project is a fully functionable agent, implemented with the A3C, as well as the PPO algorithm. Therefore, the agent needs to pass through the same training, validation and test process as referred to in [3]. The performance of said algorithms should be like the baselines in the paper, namely between 20% and 30% success rate after 5000 steps. The result will be illustrated as graphs which show the success rate over steps. Furthermore, we want to show the result of the two agents in form of a demonstration video.

**Instructions:**

- The document should not exceed two pages including the references.
- Please name the document **3DVision_Proposal_Surname1_Surname2.pdf** and upload it via the moodle.

## References

[1] P. Foo, W. H. Warren, A. Duchon, and M. J. Tarr. Do humans integrate routes into a cognitive map? map- versus landmark-based navigation of novel shortcuts. *Journal of Experimental Psychology-Learning Memory and Cognition*, 31(2):195–215, 2005.

[2] S. Gillner and H. A. Mallot. Navigation and acquisition of spatial knowledge in a virtual maze. *Journal of Cognitive Neuroscience*, 10(4):445–463, 1998.

[3] Nikolay Savinov, Alexey Dosovitskiy, and Vladlen Koltung. Semi-parametric topological memory for navigation. *ICLR*, 2018.