**Robot Exploration in Unknown Environment**

**Introduction**

Robot exploration (RE) of unknown indoor environments (UIE) has become increasingly popular due to the emergence of indoor applications of mobile robotics. (Cadena et al., 2016) RE methods are used in different (indoor) applications, such as search and rescue applications (Niroui, 2019), human assistance in crowded areas (Mammolo, 2019) and fire searching (Marjovi et al., 2021).

In the field of RE, multiple different exploration strategies exist, i.e., different paradigms, methods, and software. Though, all strategies aim for the same goal: construct a model of the UIE efficiently and/or reach a certain goal or perform a certain task (Cadena et al., 2016). Maps are useful for two reasons; (i) for path planning and to provide an intuitive visualisation for a human operator and (ii) to allow limiting the error committed in estimating the state of the robot (Cadena et al, 2016).

Moreover, the popularity of implementing reinforcement learning (RL), a subpart of machine learning (ML), has increased (Niroui et al., 2019). RL is characterised by computer learning in sequential decision-making problems in which there is limited feedback (Otterlo & Wiering, 2012). RL approaches have two limitations in RE of UIE, namely that state features need to be handcrafted and that it suffers from the curse of dimensionality (Arulkumaran, 2017). A more advanced approach of RL that is proposed is deep reinforcement learning (DRL). Contrary to RL, DRL learns state features automatically and dimensionality can be reduced through iterative interactions with the environment (Arulkumaran, 2017). Implementing DRL in RE of unknown cluttered environments, such as UIE, is still in a developing phase. (Niroui et al., 2019)

According to Cadena et al. (2016), RE is entering a new era, called *robust-perception age*, in which challenges await. These challenges are among others *robustness* and *efficiency*, which will be the two main focuses of this project. *Robustness (*or *robust performance)* entails that the RE method can operate in a broad set of UIEs with noisy (Cadena et al., 2016). Efficiency means that the robot can use less time to explore as much as possible in UIE.

/STATE OF THE ART

RE methods have developed significantly over the last decade as they have been able to operate in more complex environments for a longer period of time. (Kunze et al., 2018). /STATE OF THE ART

Due to the above-mentioned challenges and based on the state-of-the-art RE policies, a chance to design an improved RE policy based on DRL arises. Therefore, the main objective in this research is to design a RE policy based on DRL, and analyse therobustness and time-efficiency between DRL method and classical method. This policy entails that, given a robot's current pose in an UIE and the information collected by the robot from current observationso far, the goal is to produce a map of the UIE in a time-efficient manner that is as informative and as robust as possible. The robot's main task at each step is to determine the next most informative destination in the environment to move to in order to build the map.

**Stakeholder analysis**

In robot exploration two stakeholders could be distinguished, namely:

* Discrete Technology and Production Automation (DTPA); the DTPA is the main stakeholder in this project, as it is the problem owner, and therefore has high interest in this project. The DTPA has also relatively high power in this project, as they can intervene if desired.
* Overall users of RE; this stakeholder entails all people that make use of robot exploration. They could be interested in this project since a new method of robot exploration will be designed. Other users of robot exploration have very low power in this project, as they are not involved in the design of the new exploration method. However, these users do have interest in this project, as it contributes to the development of robot exploration.

No competitors are considered in this project, as most of the robot exploration libraries are open source. (Cadena et al., 2016)

**System description**

In this project, the sensor information of 2-D lidar and odometry can be seen as the observation of robot. Based on the current observation, the SLAM module is used to build environment map and calculate the location of robot in map. Based on the given map so far and the location of the robot, the robot follows a certain *exploration policy* to determine its next movement to the most informative destination. The design of this exploration policy is the main objective of this project. For the determination of a robot its next movement, there will be made use of a Markov Decision Process (MDP) as described by Otterlo & Wiering (2012). This decision will lead to a movement of the robot within the environment. Then the robot will receive new observation from the environment by its sensors and therefore a feedback loop will be created. An overview of this RE system is given in *figure 1*:

Diagram

Description automatically generated

Figure 1 - System overview of the RE method

This project will merely focus on UIE, which means that no prior knowledge is known in any way. This excludes both manually built maps of artificial beacons in the environment and GPS systems (the GPS satellites can be considered as moving beacons at known locations) (Cadena et al., 2016). Thereby, only indoor environments will be considered since this is a demarcated area and GPS has little function within indoor areas (Cadena et al., 2016).

**Research problem**

RESEARCH PROBLEM

**Goal statement**

The main goal of this design science project is to design an RE policy based on (deep) reinforcement learning. Thereafter, the designed RE policy will be compared with the traditional RE methods. Ultimately, the designed RE policy will be implemented in the real world.

These objectives result in the following goal statement:

This research aims to design and implement a robot exploration policy based on (deep) reinforcement learning which can explore unknown indoor environments time-efficiently and robustly, within a time span of 17 weeks.

This goal is considered as SMART, as it specifically aims at designing a new exploration method, measurable due to the degree of time-efficiency and robustness, attainable due to the broad variance of existing literature, relevant due to the contribution to existing RE methods, and time-bound due to the time-restriction of 17 weeks. The creation of SMART goals is based on Williams (2012).

**Main research question**

How to design and implement a robot exploration method based on reinforcement learning which explores unknown environment time-efficiently and robustly?

**Research sub-questions**

1. What causes the challenge of robustness?
   1. How to overcome the challenge of robustness?
2. What causes the challenge of time-efficiency?
   1. How to overcome the challenge of time-efficiency?
3. How to design RE policy based on reinforcement learning?
4. How does the to-be-designed RE policy compare to traditional RE policies?
5. How to implement the to-be-designed RE method in the real world?

EXPLANATION ABOUT WHY THESE QUESTION ARE RELEVANT (CHECK OTHERS)

**Research strategy**

Firstly, this research project is in-depth research, as its aim is to yield knowledge on a specific niche subject, namely robot exploration based on reinforcement learning However, the project will begin broadly by comparing the traditional methods of robot exploration.

Secondly, this research project includes both qualitative as quantitative parts. On the one hand, qualitative research is done by means of literature research about traditional robot exploration methods. Moreover, the system-thinking approach of programming the to-be-designed robot exploration methods could be considered as qualitative. On the other hand, this project will be somewhat quantitative, as the programming will include data analysis and data science.

Thirdly, the main focus and goal of this project is considered as empirical, as a robot exploration method should be designed. The implementation of this methods in the real world is even more empirical. However, in order to arrive at this stage; in the first stage desk research should be done in order to create an overview of existing methods.

**Methods and Tools**

In this project, data will be both gathered and generated. On the one hand, data about different robot exploration methods will be gathered. This includes not only information about how these methods are built up and work, but also the actual coding itself. On the other hand, data will be generated by means of designing a robot exploration method. This data generation will be based on the data gathered beforehand.

Extensive literature, modelling programmes and programming language are available and are preliminarily provided by the project its supervisor.

PYTHON ROS AND GAZEBO

**Deliverable**

The deliverable of this research project is an exploration policy based on reinforcement learning which overcomes the challenge of robustness and time-efficiency*.*

**Validation**

Internal validation will be achieved by gathering information about traditional robot exploration methods and its corresponding coding. External validation will be reached by designing an exploration method based and first simulating this method and subsequently implementing this method in a real-world environment.

**Risk analysis and feasibility**

The main risk of this research project is the time limit. Firstly, this project is rather challenging as it implies both extensive and in-depth research, which will preoccupy much time. Moreover, this project is empirical and practice oriented. Designing the robot exploration method will probably take much time.

The feasibility of reaching all the research objectives in this project is uncertain, as the time limit constraint might cause the third object, implementation in a real-world environment, to not be achieved.

**Planning**

A preliminary planning is made, which is built up in four different stages:

*Stage 1*: Orientation stage which will be used to explain and compare traditional robot exploration methods. Moreover, this stage will be used to get familiar with modelling software and programming languages. This stage will take approximately 3 to 4 weeks.

*Stage 2:* Deployment stage which will be used to get familiar with the modelling and programming of traditional exploration methods in the simulation environment. This stage will take approximately 1 to 3 weeks.

*Stage 3:* Learning-based exploration stage in which learning based methods and policies will be examined and designed. This design will be compared to traditional methods in this stage as well. This stage will take approximately 5 to 7 weeks.

*Stage 4:* Implementation stage in which the designed method will be implemented in a real-world environment. This stage will take approximately 3 to 5 weeks.

**References**

Arulkumaran, K., Deisenroth, M. P., Brundage, M., & Bharath, A. A. (2017). Deep reinforcement learning: A brief survey. IEEE Signal Processing Magazine, 34(6), 26-38.

Cadena, C., Carlone, L., Carrillo, H., Latif, Y., Scaramuzza, D., Neira, J., ... & Leonard, J. J. (2016). Past, present, and future of simultaneous localization and mapping: Toward the robust-perception age. *IEEE Transactions on robotics*, *32*(6), 1309-1332.

Keidar, M., & Kaminka, G. A. (2014). Efficient frontier detection for robot exploration. The International Journal of Robotics Research, 33(2), 215-236.

Kunze, L., Hawes, N., Duckett, T., Hanheide, M., & Krajník, T. (2018). Artificial intelligence for long-term robot autonomy: A survey. IEEE Robotics and Automation Letters, 3(4), 4023-4030.

Mammolo, D. (2019). Active slam in crowded environments (Master's thesis, Autonomous Systems Lab, ETH Zurich).

Marjovi, A., Nunes, J. G., Marques, L., & De Almeida, A. (2009, October). Multi-robot exploration and fire searching. In 2009 IEEE/RSJ International Conference on Intelligent Robots and Systems (pp. 1929-1934). IEEE.

Niroui, F., Zhang, K., Kashino, Z., & Nejat, G. (2019). Deep reinforcement learning robot for search and rescue applications: Exploration in unknown cluttered environments. IEEE Robotics and Automation Letters, 4(2), 610-617.

Otterlo, M. V., & Wiering, M. (2012). Reinforcement learning and markov decision processes. In Reinforcement learning (pp. 3-42). Springer, Berlin, Heidelberg.

Ramakrishnan, S. K., Jayaraman, D., & Grauman, K. (2021). An exploration of embodied visual exploration. *International Journal of Computer Vision*, *129*(5), 1616-1649.

Vasilii, K., Vladimir, A., Andrey, P., &amp; Vera, K. (2018). The lidar odometry in the slam. Proceedings of the Xxth Conference of Open Innovations Association Fruct, 23, 180–185. Retrieved March 24, 2022, from INSERT-MISSING-URL.

Verschuren, P., Doorewaard, H., & Mellion, M. (2010). *Designing a research project* (Vol. 2). The Hague: Eleven International Publishing.

Williams, C. (2012). MGMT (5th ed.). USA: South- Western College Publishing.