Summary of the Meetings

27.7

First Meeting

- present ideas of the project
- discussing scope of ROS and talking about alternative programs
- QUESTION: how do we want to organize the project (time/ how to divide workload)
- HOW to measure reward? (**driving against walls** / others difficulties -- look for more papers)

  → innovation metric comparing produced maps

- computational resources for this problem?
- ALTERNATIVE? Slam applied to different environments
- find a new meeting date? Regular dates?

TODO:

- discuss the workflow we want to have (DEADLINE 18.9)
  - Ziel : 31. August fertig mit Project
  - 2-3 weekly meeting with group to update
  - 1/2 mal die Woche mit Leon meeting
  - Github; Aufgabenteilung
  - (UML-) Diagramm
- look at simulators → choose one
  - too complex, the simpler ones we found were not well documented

  → we won't use simulators

- read papers to figure out a good reward system
  - https://ieeexplore.ieee.org/document/9238090 if we implement the complex version
  - https://arc.aiaa.org/doi/abs/10.2514/6.2019-0396 if we implement the simplified version in a self made grid world
- maybe: find a toy problem - for example gridworld to simplify project
  - create own gridworld
- find date for next meeting(s)
  - 11./ 12. / 13. August 14:30 - 18

Further ideas:

+ continuously negative reward learns better than just rewarding findings in the environment ?

2nd meeting 13.08

- gridworld may be a bit simple → boring?
    - → can add more complexity
        - – e.g. slam in the gridworld/ measurement noise
        - test different reward systems (negative reward for no information gain)
- https://arxiv.org/pdf/2012.02096.pdf
- https://arxiv.org/abs/2103.01991
    - maybe use some gridworld generating approaches from this paper

Third Meeting 13.9

to discuss:

- update what we've done so far / what do you think
- scope of the project
- Reward values?
    - generally the values are not as important as their relation
    - try to keep the values between 0 and 1 in the end
    - makes sense to make the negative reward the dominant one, for example keep the positive one smaller that it by a factor of 10
        - so: negative reward each time step and positive reward for each discovered tile - with the positive reward functioning as a shaping reward
        - dense reward problem
- video:
    - 5 to 10 minute length
    - short description of the project
        - what is the problem
        - why is this difficult
        - how did we solve it
    - don't go too much into details, assume general knowledge of RL

    - think of it as short showcase video that can be shown in the repo to show people what can be done at the end of the course and to introduce the project
- umfangreiches Paper? / Video: Wie genau soll das werden?
    - was haben wir gemacht um das Problem zu lösen quasi für leihen projekt vorstellen max 5-10 Minuten
    - scientific project report; CITATION / wissenschaftliche Sprache;

→ journey/videos von performance

Group meetings

meeting: 15.08. 15:00

TODO for that meeting:

- read paper about gridworld generation
- [x] ~~Code von paper zum laufen bekommen PAIRED~~
- [ ] anpassen an needs von unserem Projekt
    - [x] ~~kein Goal~~
    - [ ] Parameter?
    - [x] ~~Visualisieren der grid worlds~~
- [ ] Reward?
    - [ ] Was wenn der Raum z.B Kreis, aber ein Grid wird erstellt

        → funktion, die erkennt, dass form abgeschlossen -- finished map indicator

Friday, 26 August, 2022

TODO till then:

- search for usage of AdversarialEnv
- → if nothing useful is found, have meeting with Leon
- write subclass for multigrid to replace AdversarialEnv

DONE:

- implemented render method and figured out which env variables hold objects of which class
- started implementing class to replace AdversarialEnv with custom step function, reward, and including the map

To implement:

- reward: small negative reward each time step, positive reward for each new piece of information

meeting: 28.08. 15:00

TODO till then:

- find values for reward → find examples from papers
- finish implementing custom class with map and reward → Hannah
- think about map implementation — how to make it more general/how to deal with having less information, e. g. no defined grid size, no location information
- Agent Code übernehmen? → Lisa

NEXT MEETING:

TODO till then:

- look for papers about finding closed contours
    - map is completed when there is no tile that the agent can walk on that connects to an unknown tile → it should suffice to look for such tiles instead of finding an enclosing contour
    - this could be checked when the map is updated, as that is when the status of the newly updated tiles and their neighbors changes to
- Argumentation wie wir policy setzen (wie im Paper)
- find values for reward → find examples from papers

Meeting 11.9.2022

- Parametervergleiche -- Einfluss auf die Performance?
- github repo
- organization for project (time management: video, paper, …)
- start with paper and structure
- map implementation
- Umsetzung: Testen von noisy images? vorerst auslassen, zu zeitintensiv

Monday, 12 September, 2022

- flexible paired version already implemented :(
- look at implementing budget paired → requires us to find something comparable to the SKIP action they used in the paper
- Was wollen wir testen? (rewards, grid worlds, …)
- how to record results? kurze Video Clips, wie Agent performt?
- Was wollen wir alles haben und was ist bereits davon schon implementiert?

Wednesday, 14 September, 2022

- budget Paired Paper angesehen und über Implementierung nachgedacht
  - lösbar denken wir
- Paper Motivation und Introduction (Vergleich DRL / non-DRL approaches)
- Gedanken für die Planung / Umsetzung
- Was wollen wir testen? - Wie lange dauert ausführung und davon ausgehen Entscheidung

Plan A : Leon antwortet bis morgen

+ er sagt der Code reicht, Daumen hoch
- Er sagt, Code reicht nicht --- können wir das realistisch noch ändern? Abwägen

Plan B : Leon meldet sich nicht, oder erst (Nach) Freitag?!

- Deadline Extension? Starker Grund, nicht gemeldet?

Time planning

•finish implementation until Thursday evening

• run training on Friday

• continue writing paper on Friday

• finish writing paper (results and visualization) and video on the weekend