JISDM 2022
5th Joint International Symposium on Deformation Monitoring
20 -22 June 2022
Polytechnic City of Innovation, Valencia, Spain jisdm2022.webs.upv.es

Adaptive spatial discretization using reinforcement learning

Jemil Avers Butt[1,2], Andreas Wieser[1]
[1]ETH Zurich, Institute of Geodesy and Photogrammetry, Zurich; [2]Atlas optimization GmbH, Zurich

ETH zürich

ATLAS OPTIMIZATION

UNIVERSITAT POLITÈCNICA DE VALÈNCIA

International Federation of Surveyors

International Association of Geodesy

isprs information from imagery

Content

1. Problem

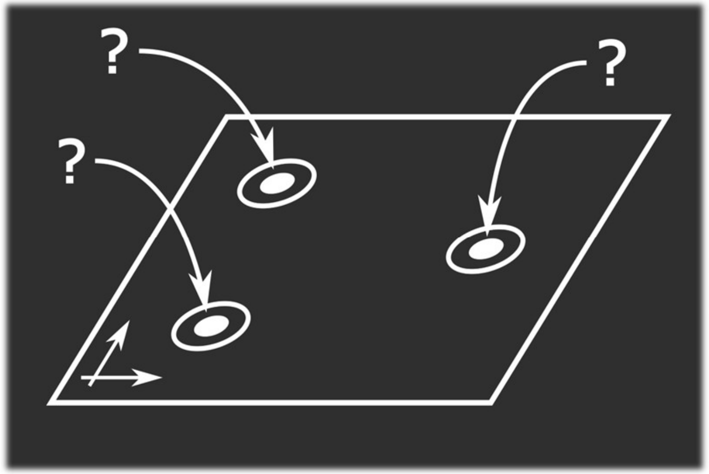
Spatial discretization

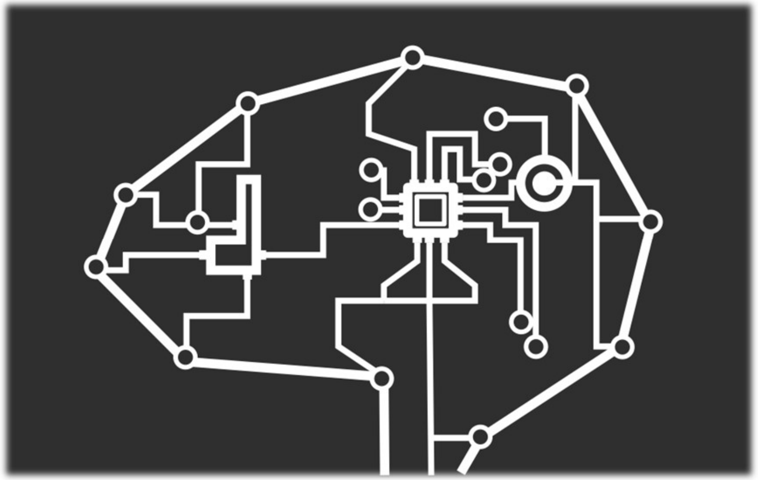
2. Method

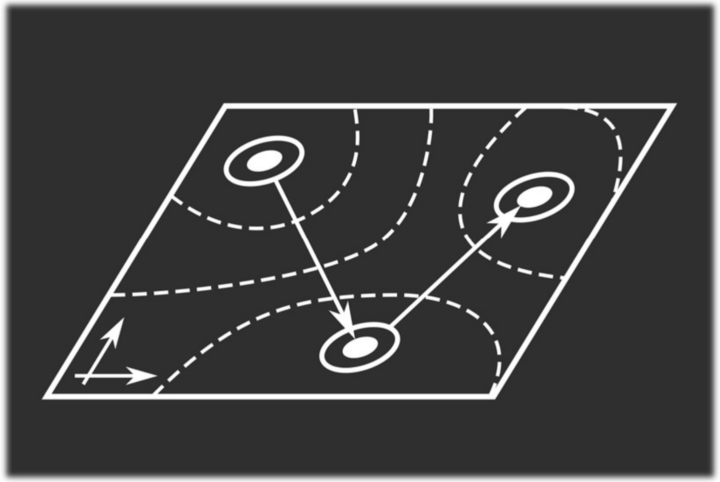
Reinforcement learning

3. Results

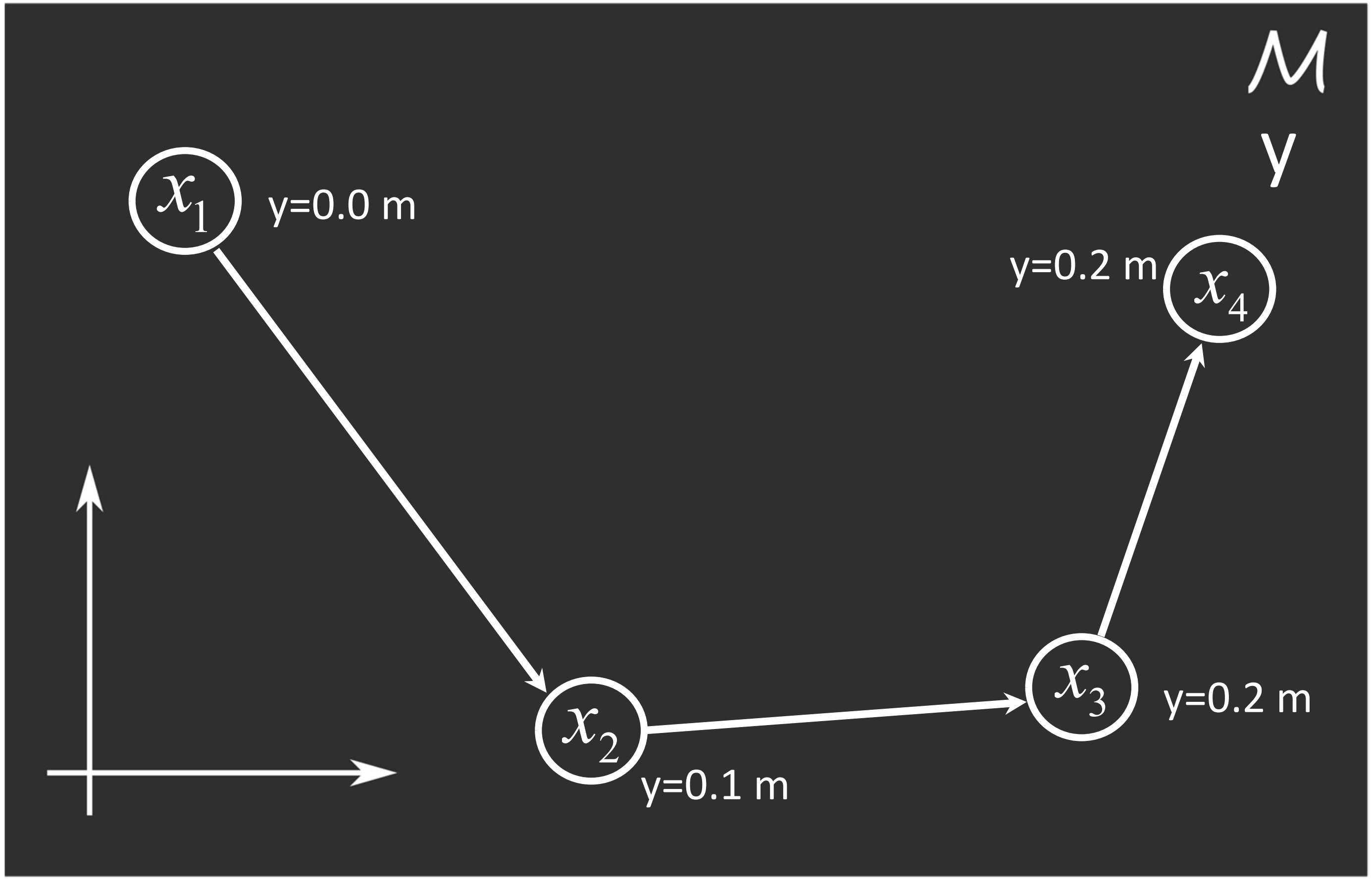
Better sampling

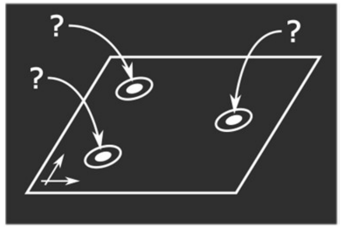






Problem - Specification





Spatial discretization

- Object M
- Interested in y
- Sequence of $x_1, x_2, \ldots$

Questions

- Best sequence?
- System knowledge?
- Adapt to observations?

Answer

- We need a policy π !

Problem – Hierarchies of Solutions

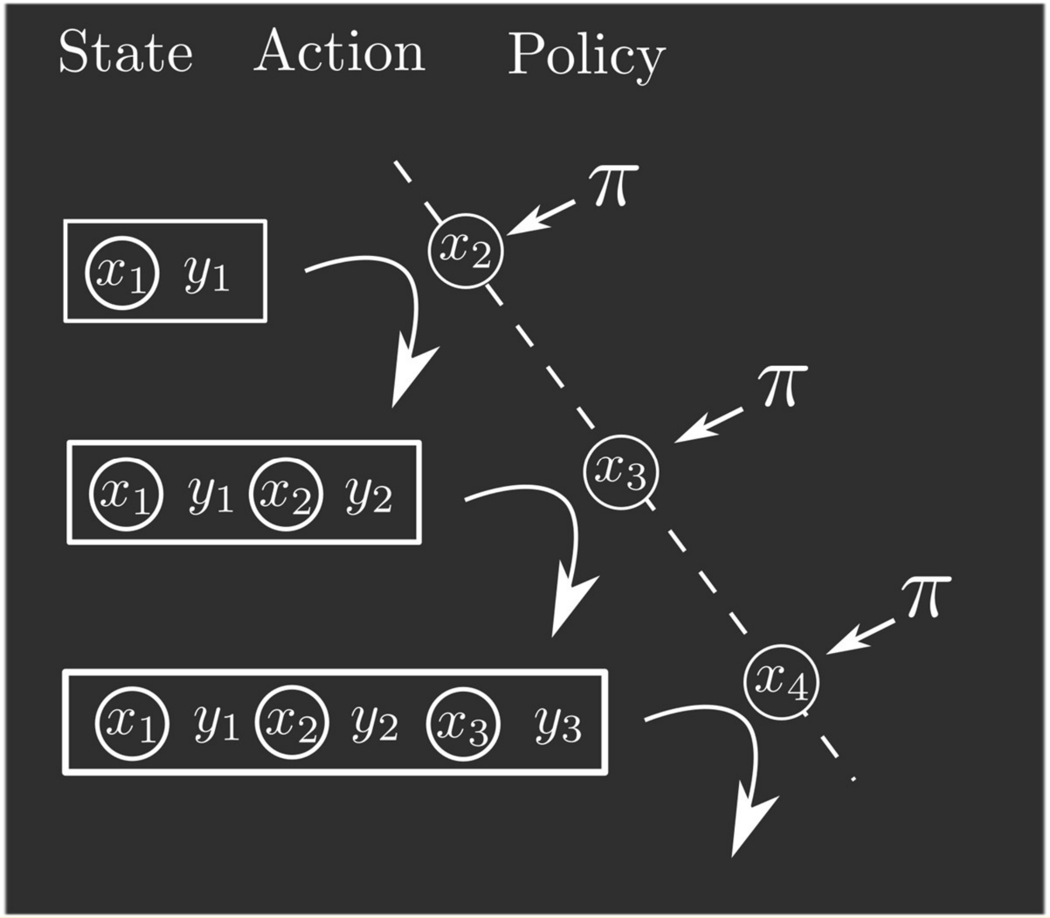
$\pi(\cdot)$ is called policy

π : state $\mapsto$ action

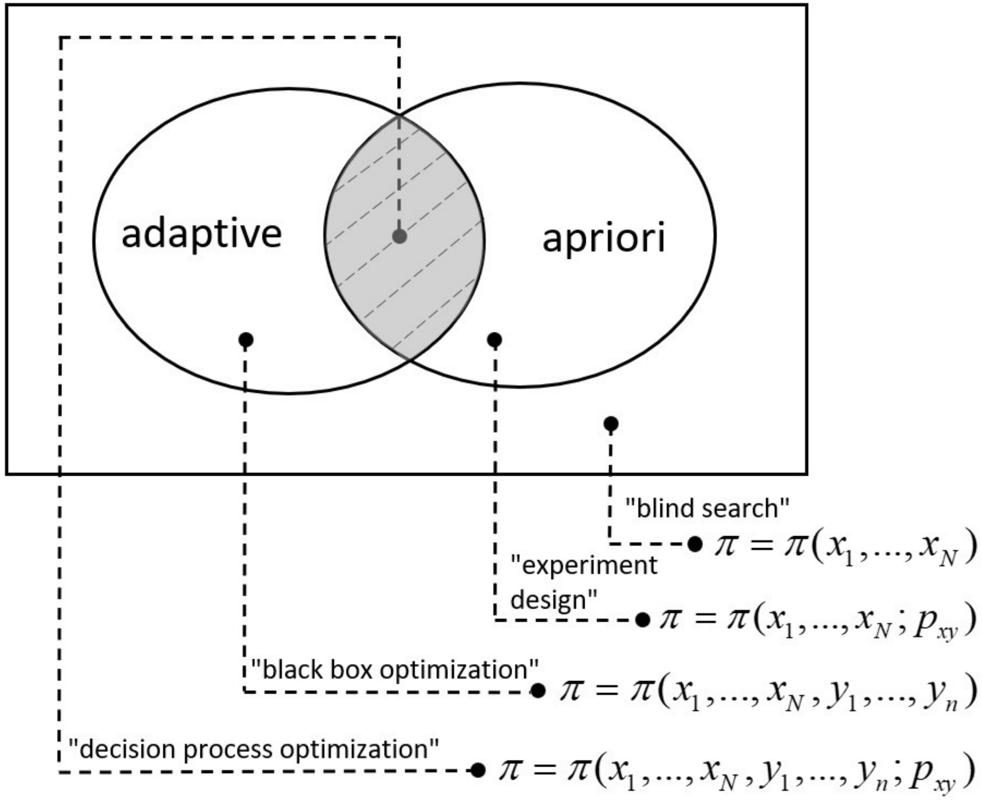
What we know — Next measurement

We want

- Adaptivity
  Next sample locations influenced by pevious findings
  $\rightarrow \pi = f(x_1, \ldots, x_N, y_1, \ldots, y_N)$
- Prior knowledge
  Sampling scheme makes use of dependences between x, y
  $\rightarrow \pi = f(x_1, \ldots, x_N; p_{xy})$



Problem – State of the art



"Blind search" :

Random numbers, pseudorandom, grids, ...

[Kuipers & Niederreiter]

"Experiment design" :

Minimization of $\mathrm{tr}(\Sigma)$, $\det(\Sigma)$, Entropy,...

- > Geodetic networks, geostatistical sampling

[Grafarend & Sanso], [Angulo et al]

"Black box optimization" :

Sequences of samples based assumptions.

-> Problems with unknown structure

[Fu], [Alarie et al]

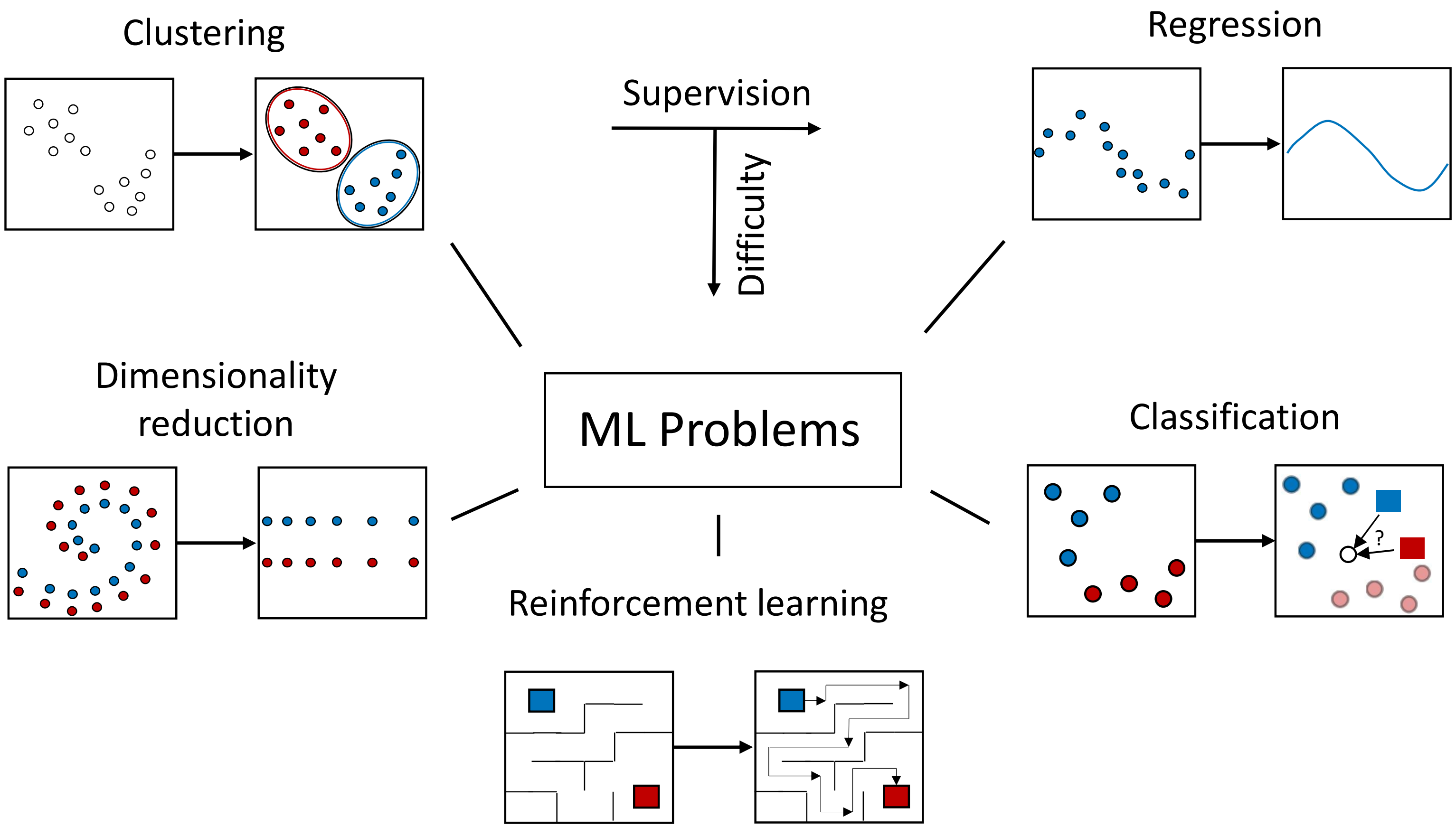
"Decision process optimization" :

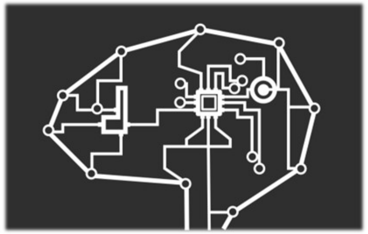
Adaptive policies to maximize reward

-> Optimal control in economy and games

[Sutton & Barto],
[Feinberg & Shwartz]

Method – RL in ML





RL:

- Optimal control
- No right/wrong
- Learn from experience
- Agent interacts with environment
- Difficult task
- Computationally intense

Method – RL Policy gradients

Policy function is ANN
- Determines decisions
- Make decisions optimal

1. Expected return depends on params

$$\eta(\pi) = E_{\pi p}\left[\sum_{k=1}^{n-1} \gamma^k r(s_k, a_k)\right]$$

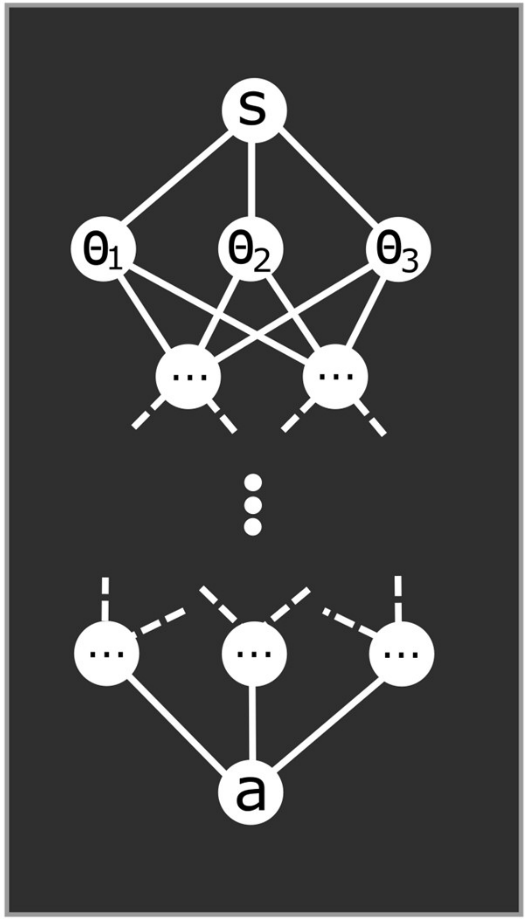
$$\eta(\pi) = \int_{(S\times A)^{n+1}} \sum_{k=0}^{n-1} \gamma^k r(s_k, a_k) p_\theta(\xi) d\xi$$

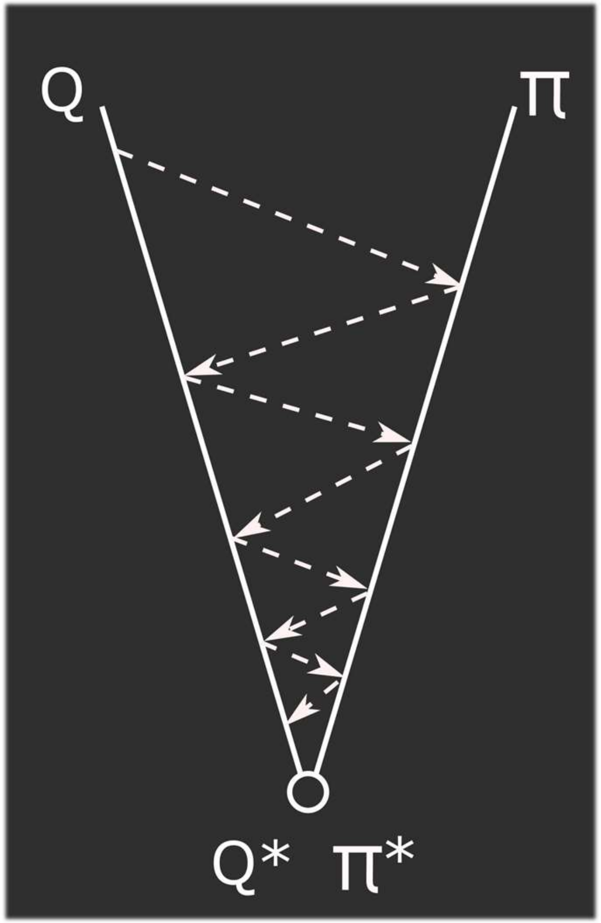
2. Gradient depends on observations

$$\nabla_\theta \eta = \int_{(S\times A)^{n+1}} \sum_{k=0}^{n-1} \gamma^k r(s_k, a_k) \nabla_\theta [\log p_\theta(\xi)] p_\theta(\xi) d\xi$$

$$\hat{\nabla}_\theta \eta = \frac{1}{m}\sum_{j=1}^{m}\left(\sum_{k=0}^{n-1} \gamma^k r(s_k^j, a_k^j)\right) \nabla_\theta [\log p_\theta(\xi_j)]$$

$$= \frac{1}{m}\sum_{j=1}^{m}\left(\sum_{k=0}^{n-1} \gamma^k r(s_k^j, a_k^j)\right) \sum_{i=0}^{n-1} \nabla_\theta \log \pi_\theta (s_i^j, a_i^j).$$



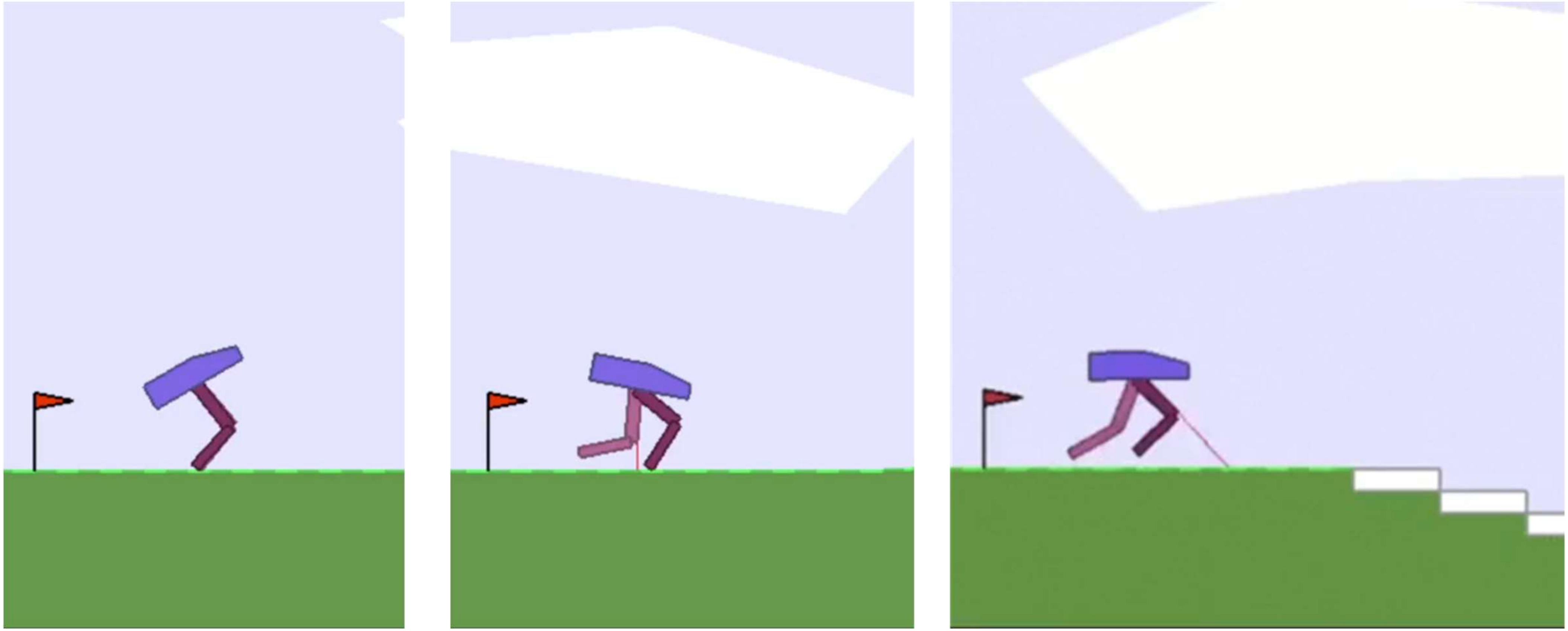


Method – Showcase

Training: 3 min

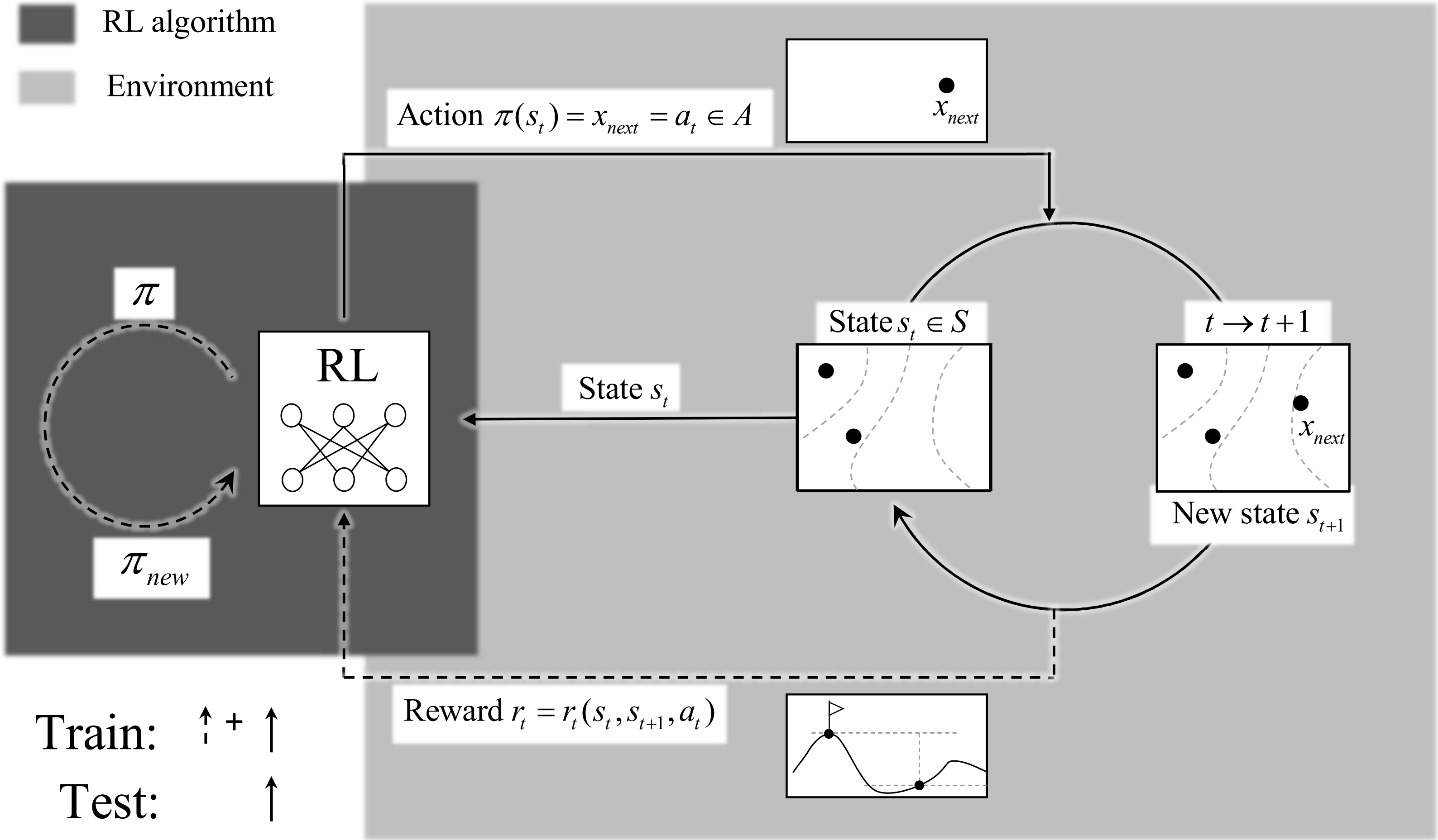
Training: 3 h

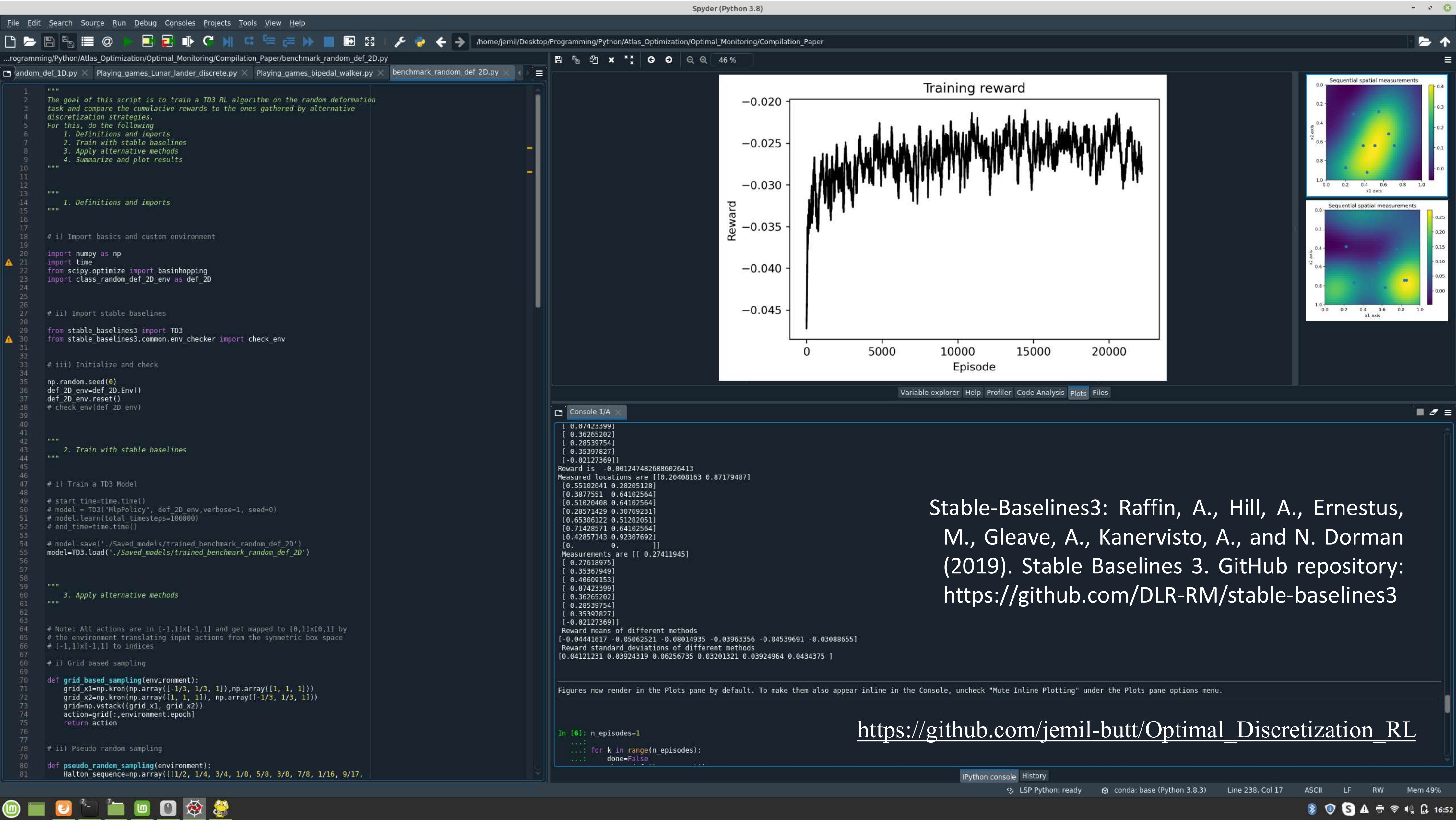
Training: 3 d



Trained with a3c by David Griffis
https://github.com/dgriff777/a3c_continuous

Method – Spatial discretization as RL

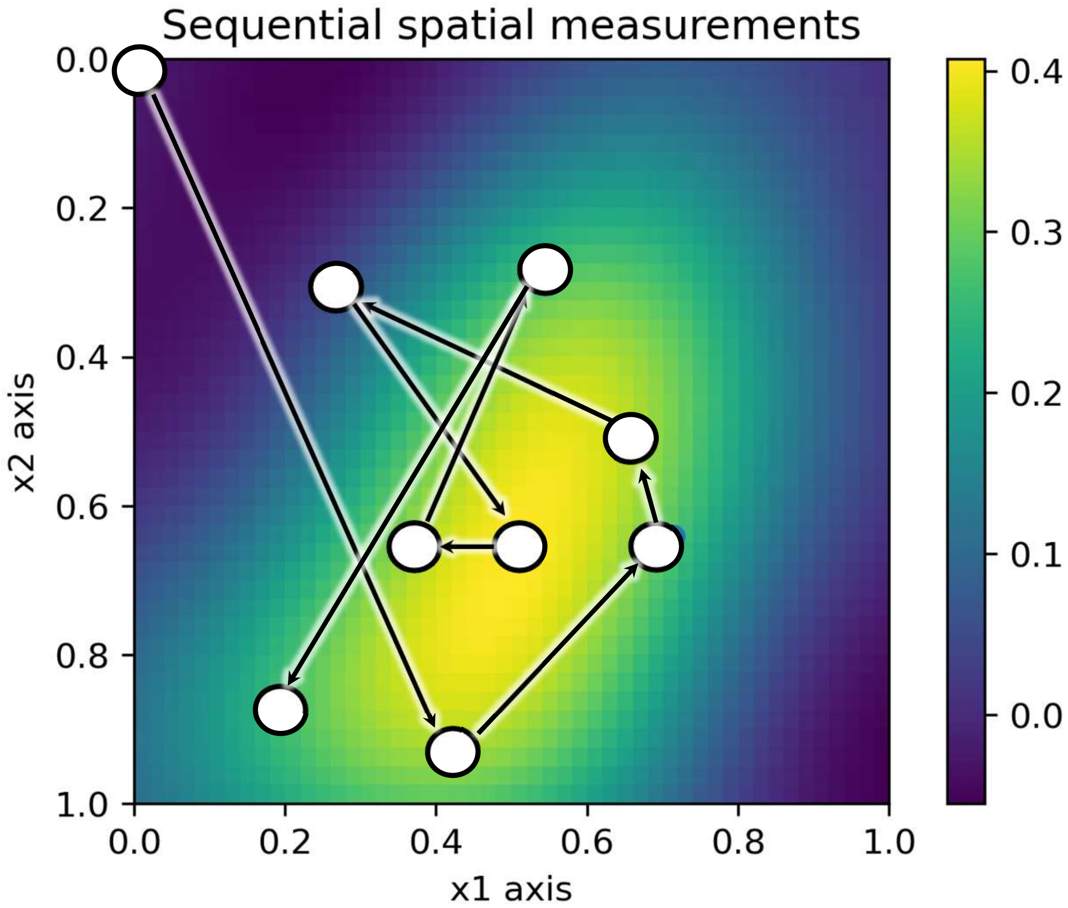




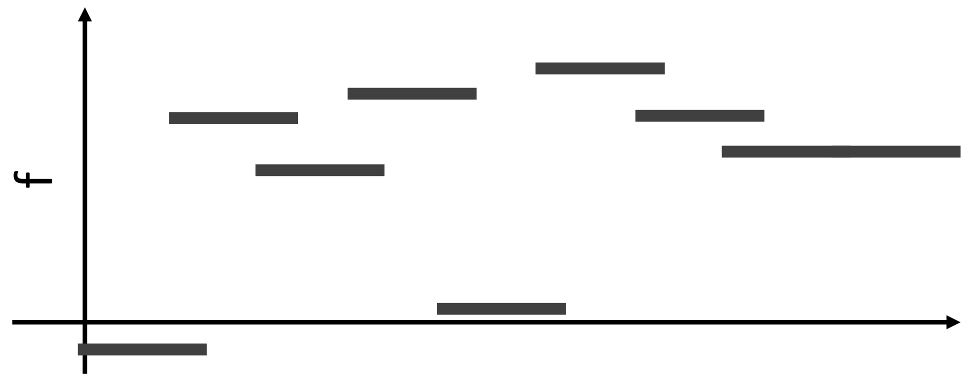
Stable-Baselines3: Raffin, A., Hill, A., Ernestus, M., Gleave, A., Kanervisto, A., and N. Dorman (2019). Stable Baselines 3. GitHub repository: https://github.com/DLR-RM/stable-baselines3

https://github.com/jemil-butt/Optimal_Discretization_RL

Results – Illustration



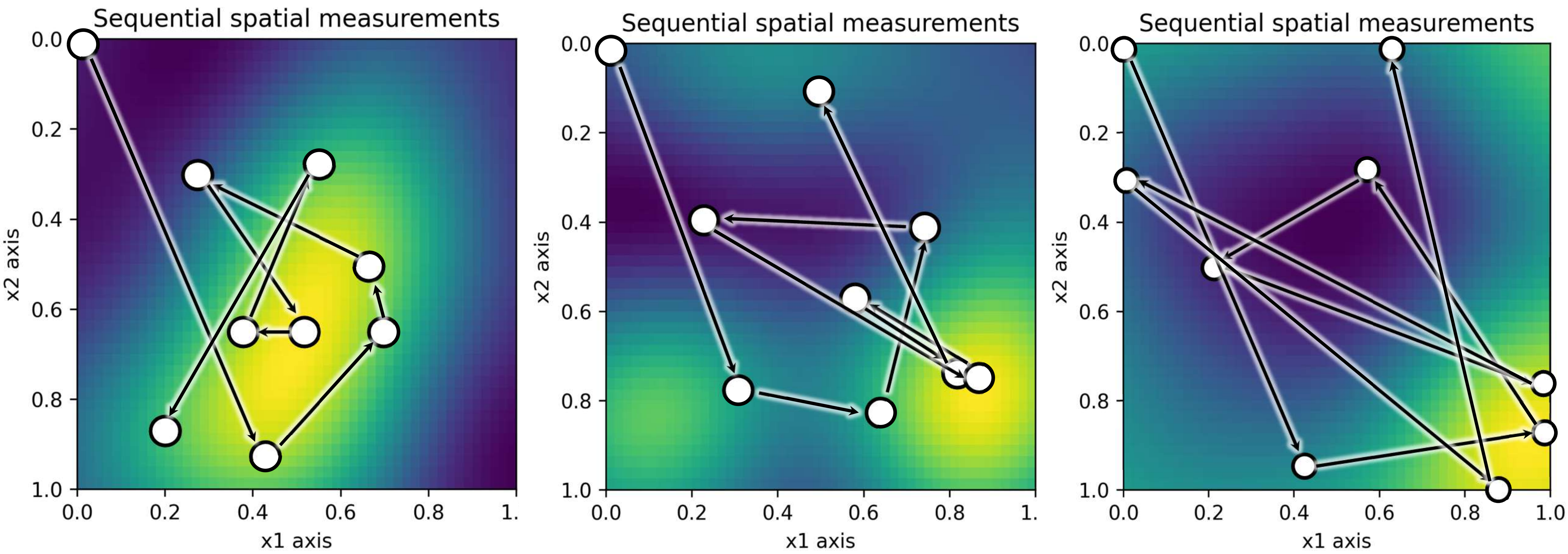
Meas nr. 1 2 3 4 5 6 7 8 9



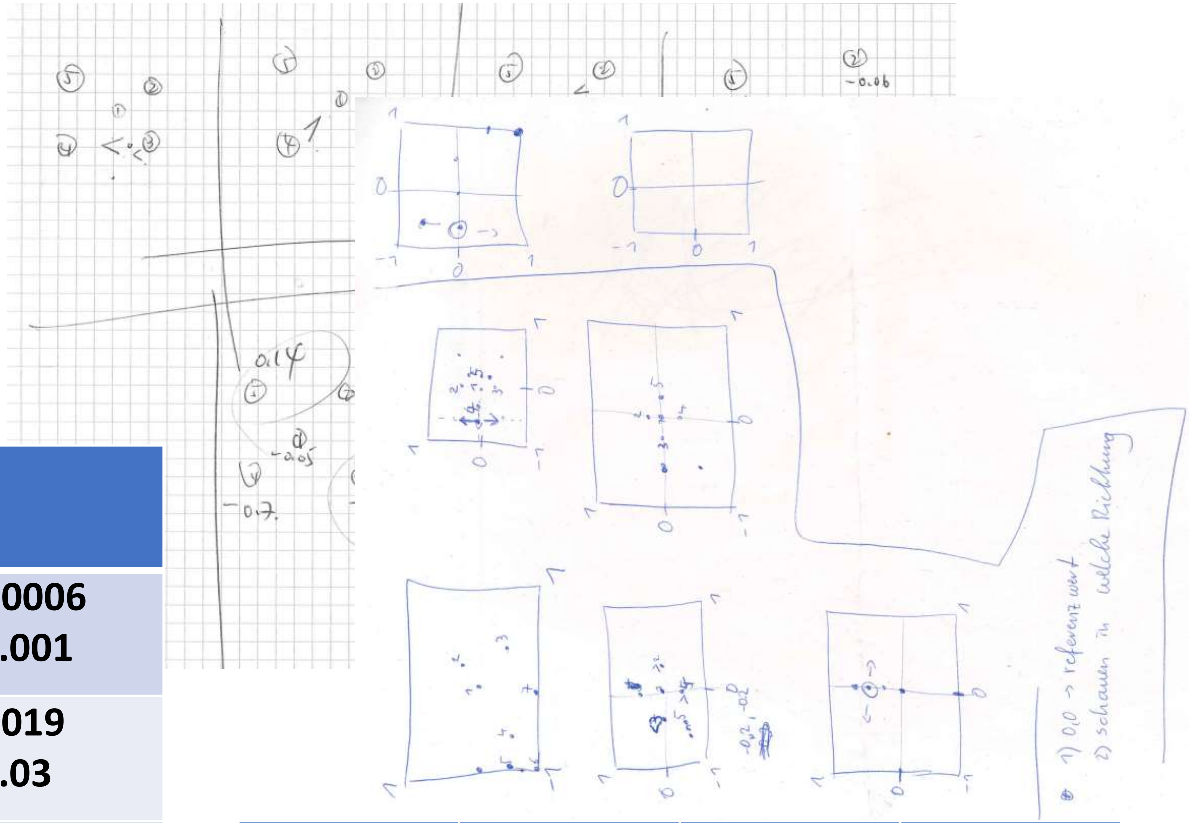
f $[-0.02]$ $[0.35]$ $[0.28]$ $[0.36]$ $[0.07]$ $[0.40]$ $[0.35]$ $[0.27]$ $[0.27]$

Locations $\begin{bmatrix}0\\0\end{bmatrix}$ $\begin{bmatrix}0.42\\0.92\end{bmatrix}$ $\begin{bmatrix}0.71\\0.64\end{bmatrix}$ $\begin{bmatrix}0.65\\0.51\end{bmatrix}$ $\begin{bmatrix}0.28\\0.30\end{bmatrix}$ $\begin{bmatrix}0.51\\0.64\end{bmatrix}$ $\begin{bmatrix}0.38\\0.64\end{bmatrix}$ $\begin{bmatrix}0.55\\0.28\end{bmatrix}$ $\begin{bmatrix}0.20\\0.87\end{bmatrix}$

Results – Illustration



Results – Comparison



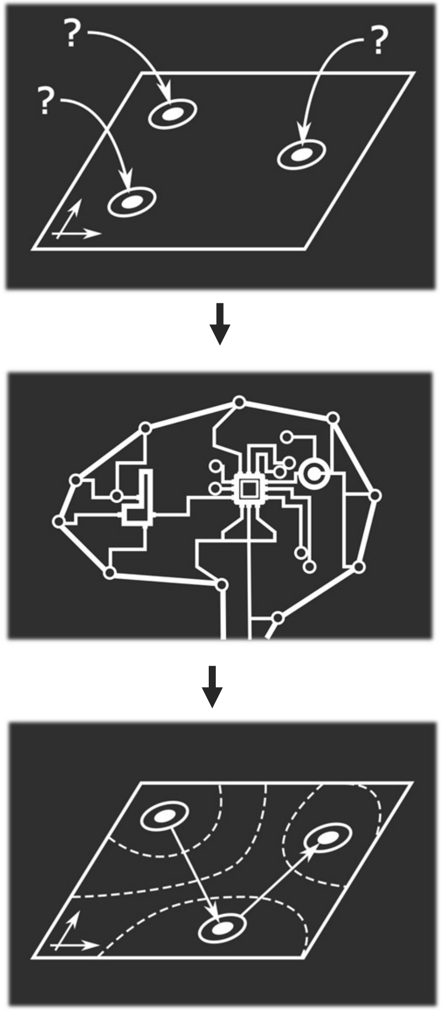
Method/ Task	Random	Quadrature	Exp. design	RL
beam bending	-0.05 ±0.1	-0.065 ±0.07	-0.0025 ±0.005	**-0.0006 ±0.001**
1D def.	-0.07 ±0.08	-0.047 ±0.04	-0.035 ±0.04	**-0.019 ±0.03**
Def. tracking	-0.28 ±0.19	N. A	-0.24 ±0.14	**-0.067 ±0.038**
2D def.	-0.08 ±0.063	-0.04 ±0.032	-0.045 ±0.039	**-0.031 ±0.043**

Human*	Human 1	Human 2	Human 3
2D def.	-0.078	-0.015	-0.026

Results – Conclusion

Problem: Have some object or process. Want to know something. Decide where to measure.

Solution: Make model of process. Use it for simulations. RL gets optimal* sequences of decisions.



What's good!

- Really flexible
- Optimal decisions under uncertainty
- Don’t need explicit probabilistic models
- Modelling work done by computer
- Higher score than alternatives

What's bad!

- No guarantees
- No measure of uncertainty
- Need model for simulation
- Training sloooow …

Applications!

- Adaptive monitoring
- Episodic repositioning

Sources

Ai, B., Jia, M., Xu, H., Xu, J., Wen, Z., Li, B. and D. Zhang (2021). Coverage path planning for maritime search and rescue using reinforcement learning, Ocean Engineering, Volume 241.

Alarie, S., Audet, C., Gheribi, A.E., Kokkolaras, M., and S. Le Digabel. (2021). Two decades of blackbox optimization applications, EURO Journal on Computational Optimization, 9.

Angulo, J.M., Ruiz-Medina, M.D., Alonso, F.J. and M.C. Bueso (2005). Generalized approaches to spatial sampling design. Environmetrics, Vol 16, No. 5.

Feinberg, E. A. and A. Shwartz (2012). Handbook of Markov Decision Processes. Methods and Applications. Berlin Heidelberg (Springer Science & Business Media).

Fu, M. C. Handbook of Simulation Optimization. Berlin, Heidelberg: Springer, 2014.

Grafarend, Erik W. and F. Sansò (2012). Optimization and Design of Geodetic Networks. Berlin Heidelberg (Springer Science & Business Media).

Sutton, R.S. and A.G. Barto (1992). Reinforcement Learning. Berlin Heidelberg (Springer Science & Business Media).

Wiering, M. and M. Otterlo (2012). Reinforcement Learning. State-of-the-Art. Berlin Heidelberg (Springer Science & Business Media).

Zuluaga, J. G. C., Leidig, J. P., Trefftz, C., and G. Wolffe, (2021). Deep Reinforcement Learning for Autonomous Search and Rescue, *NAECON 2018 - IEEE National Aerospace and Electronics Conference*, pp. 521-524.

Thank you for your time!

Now:

Q & A