IDALAB

PARIS
LODRON
UNIVERSITÄT
SALZBURG

Simon Hirländer

RL Coffee

Is RL the right tool?

- Optimisers:

➡ Always re-explore - no memory → RL can

➡ Cannot handle delayed consequences → RL can

- Accelerators seem to be generally a good environment for RL:

➡ Known reward e.g. Intensity (nevertheless hard to design)

➡ The state defined though beam diagnostics

➡ The actions are mostly well designed

- Open issues:

➡ Is a sufficient state available?

➡ How to deal with non-stationarity?

➡ How to improve the sample efficiency?

➡ Stability - how to tune the algorithms?

➡ What about safety?

66

Is RL the right tool?

- Optimisers:
  - ➡ Always re-explore - no memory → RL can
  - ➡ Cannot handle delayed consequences → RL can
- Accelerators seem to be generally a good environment for RL:
  - ➡ Known reward e.g. Intensity (nevertheless hard to design)
  - ➡ The state defined though beam diagnostics
  - ➡ The actions are mostly well designed
- Open issues:
  - ➡ Is a sufficient state available?
  - ➡ How to deal with non-stationarity?
  - ➡ How to improve the sample efficiency?
  - ➡ Stability - how to tune the algorithms?
  - ➡ What about safety?

Some directions

- Observability (POMDPs)
- Better diagnostics
  Extend state formulation (e.g. longer history) Complication: additional prediction problem
- Sample efficiency
- Develop more sample efficient algorithms - e.g. model-based Safety
- Modelling - tricky in stochastic settings (accurate probabilistic models)
- Non-stationarity (special POMDPs)
- Everyday: short horizon problems - classical and Bayesian optimisation Again: extend state formulation