





# IDALAB

EFFICIENT DATA ANALYTICS SOLUTIONS



PARIS  
LODRON  
UNIVERSITÄT  
SALZBURG



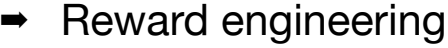
FOR AIRPLANE

# Challenges of RL in the real world

- Problem formulation - capturing the problem in an Markov Decision Problem (MDP)

State representation, Markov Property (e.g. non stationarity)





Reverse engineering



• **RL-Goals:**

sanpiedi etficiency



stability



Hyperparameter tuning



Expansion





startv

**Robustness to Changes**

→ Generation



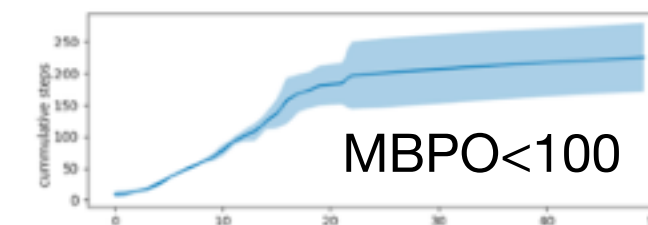
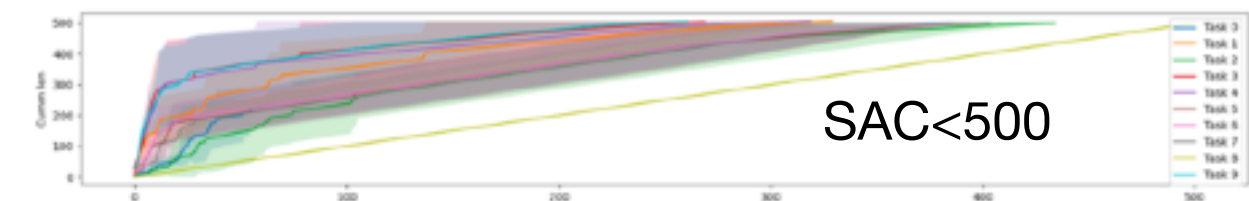
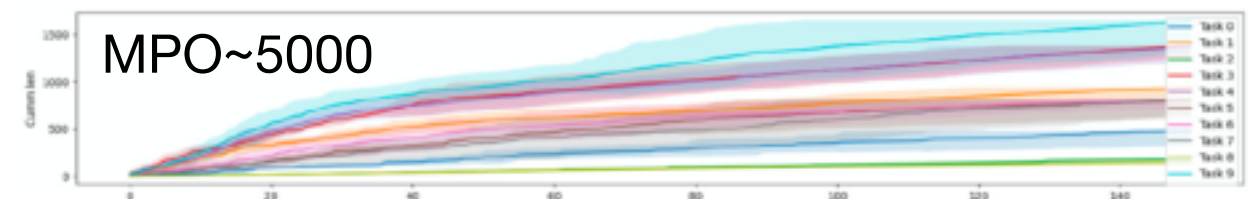
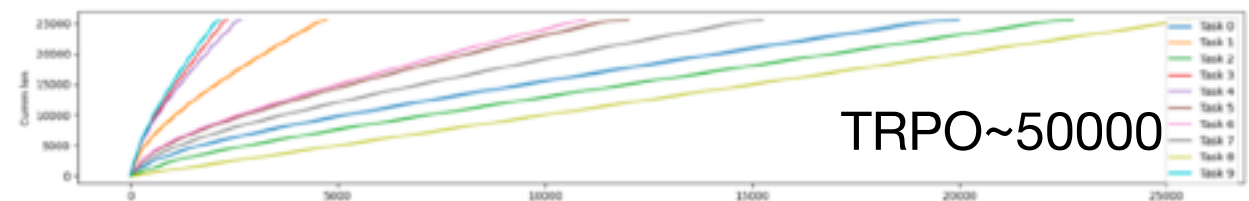


# Challenges of RL in the real world

- Problem formulation - capturing the problem in an Markov Decision Problem (MDP)
  - ➔ State representation, Markov Property (e.g. non stationarity)
  - ➔ Reward engineering
- RL - core issues:
  - ➔ **Sample efficiency**
  - ➔ **Stability**
  - ➔ **Run time**
  - ➔ **Hyper-parameter tuning**
  - ➔ **Exploration**
  - ➔ **Safety**
  - ➔ **Robustness to Changes**
  - ➔ **Generalisation**
  - ➔ ...

# Sample efficiency

- Derivative free methods: (NES, CMA,...)
- 10 x Online methods (A3C)
- 10 x Policy-gradient methods (TRPO)
- 10 x Replay-Buffer + Value function estimation (Q-Learning, DDPG, TD3, NAF, SAC,...)
- 10 x Model-based RL methods (MPO, Guided Policy Search, Dyna)
- 10 x Model-based shallow methods (no NNs) Few shot GPs...



GP-MPC < 50