IDALAB

PARIS
LODRON
UNIVERSITÄT
SALZBURG

Simon Hirländer

Tutorial RL4AA

POMDPs and non stationarity

- To find a proper state we have to solve the additional prediction problem

$$s_t = h_t(o_t, a_{t-1}, o_{t-1}, a_{t-2}, o_{t-2} \dots)$$

- In the non-stationary, finite horizon formulation the MDP has the form $(S, A, \{P\}_h, \{r\}_h, H, \rho_0)$ $\Rightarrow$ Value-functions $Q_h(s, a)$ get time depended
  $\Rightarrow$ similar form of Bellman equations

- We can incorporate time into state e.g. $\tilde{s} = (s, h) \Rightarrow$ standard MDP

- Generally Bellman equation nice in discounted, stationary formulation ⇒ this is what we usually see and most libraries build on this formulation

16

POMDPs and non stationarity

- To find a proper state we have to solve the additional prediction problem
  $s_t = h_t(o_t, a_{t-1}, o_{t-1}, a_{t-2}, o_{t-2} \dots)$
- In the non-stationary, finite horizon formulation the MDP has the form $(S, A, \{P\}_h, \{r\}_h, H, \rho_0) \Rightarrow$ Value-functions $Q_h(s, a)$ get time depended
  $\Rightarrow$ similar form of Bellman equations
- We can incorporate time into state e.g. $\tilde{s} = (s, h) \Rightarrow$ standard MDP
- Generally Bellman equation nice in discounted, stationary formulation $\Rightarrow$ this is what we usually see and most libraries build on this formulation

The entire problem