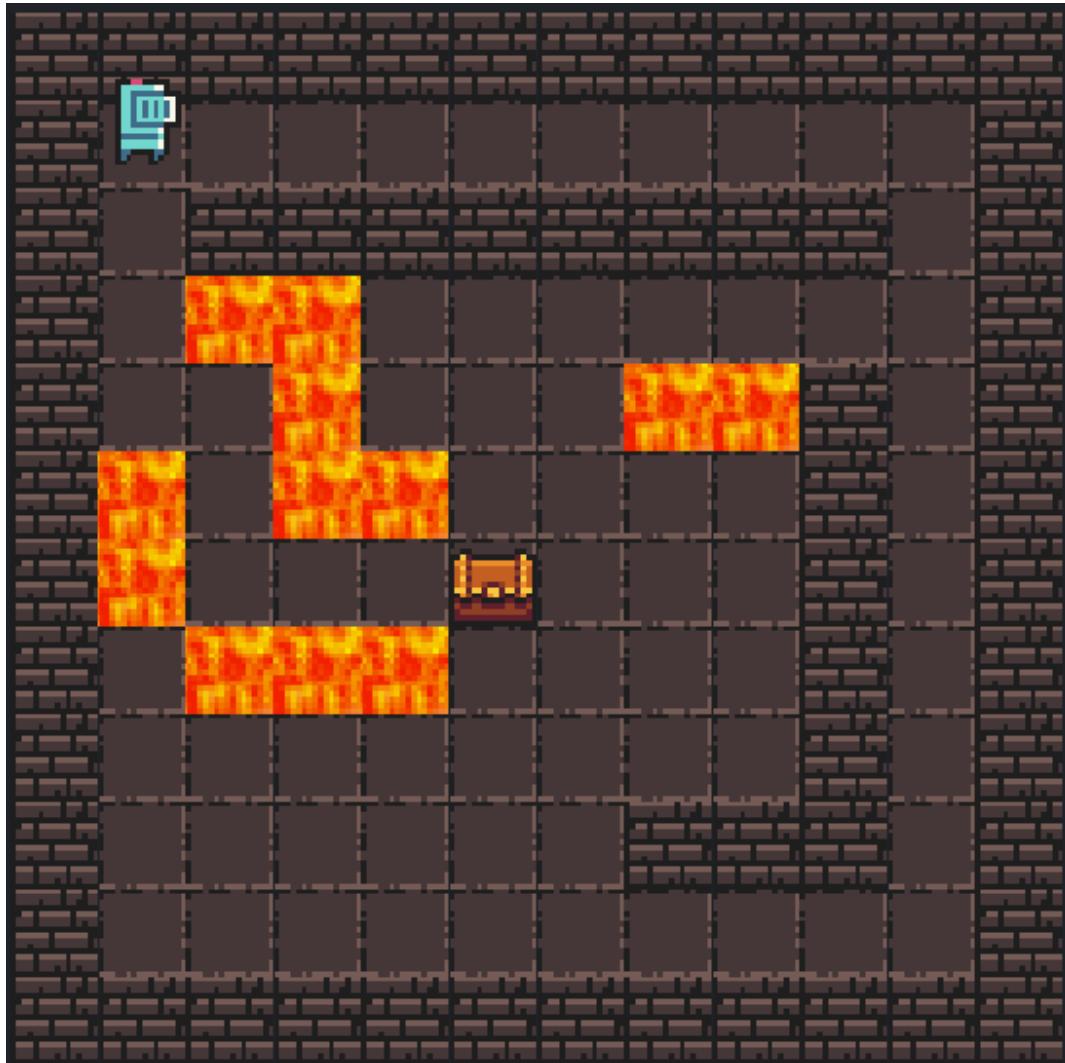


Peril, Prudence and Planning as Risk, Avoidance and Worry

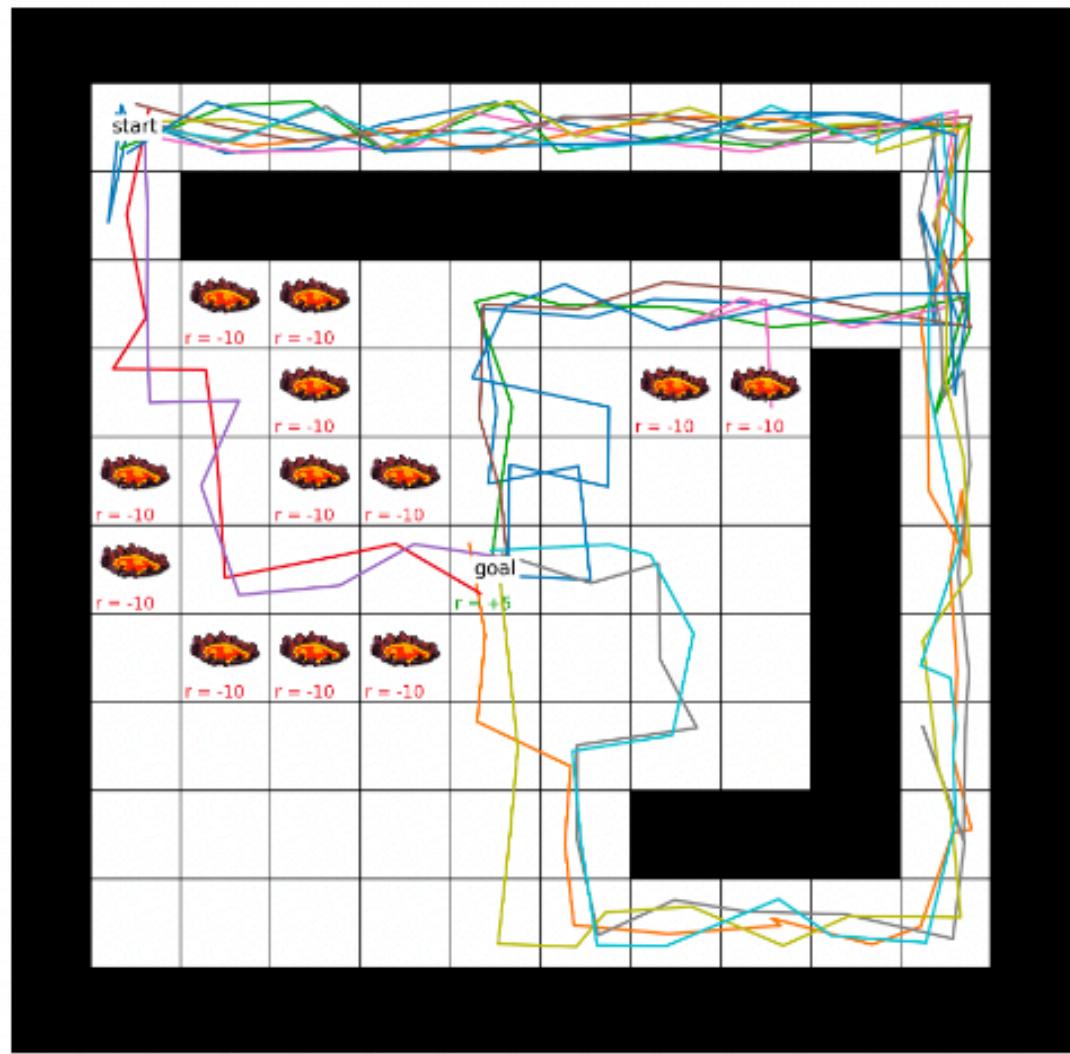
Peter Dayan

MPI for Biological Cybernetics, University of Tübingen

Chris Gagne Yannick Streicher



- -10 for lava
- +5 for reward
- 12.5% error
- 0.9 discount factor



Plan

- risk aversion
- conditional value at risk CVaR in sequential problems
 - pre-committed pCVaR
 - nested nCVaR
- risk averse on-line behaviour
- risk averse off-line planning
 - replay and rumination

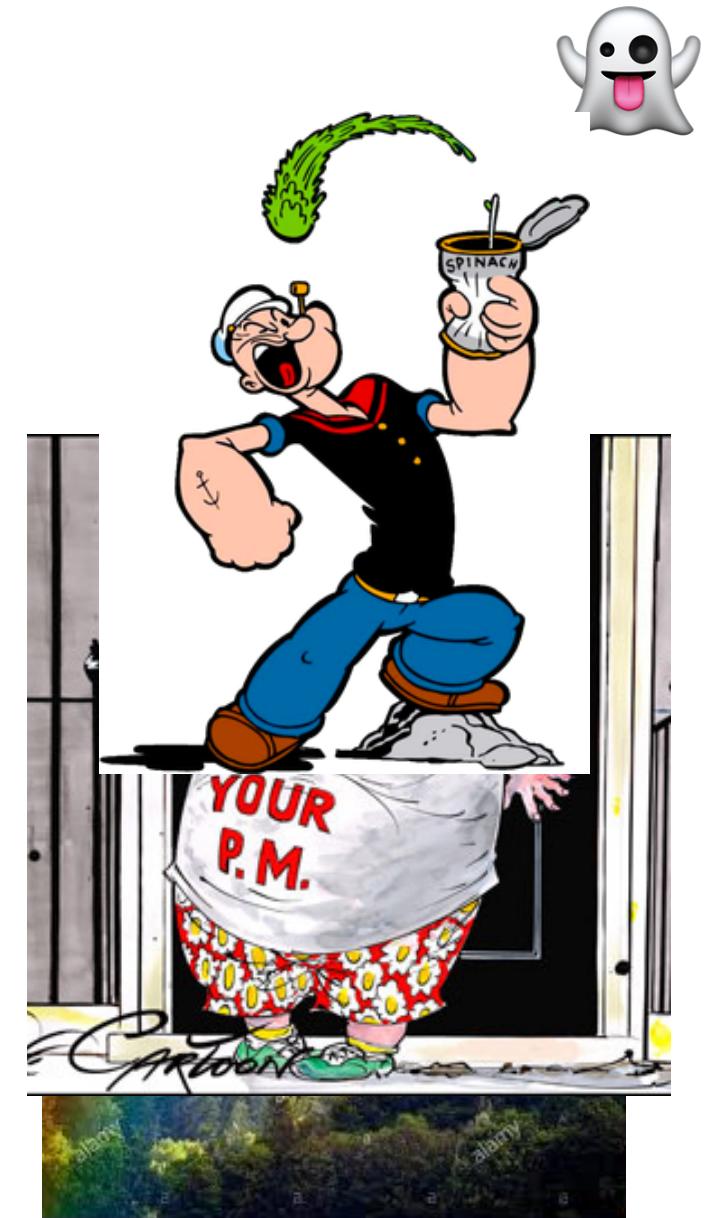
the wrong environment

Computational psychiatry

the wrong problem



the wrong solution



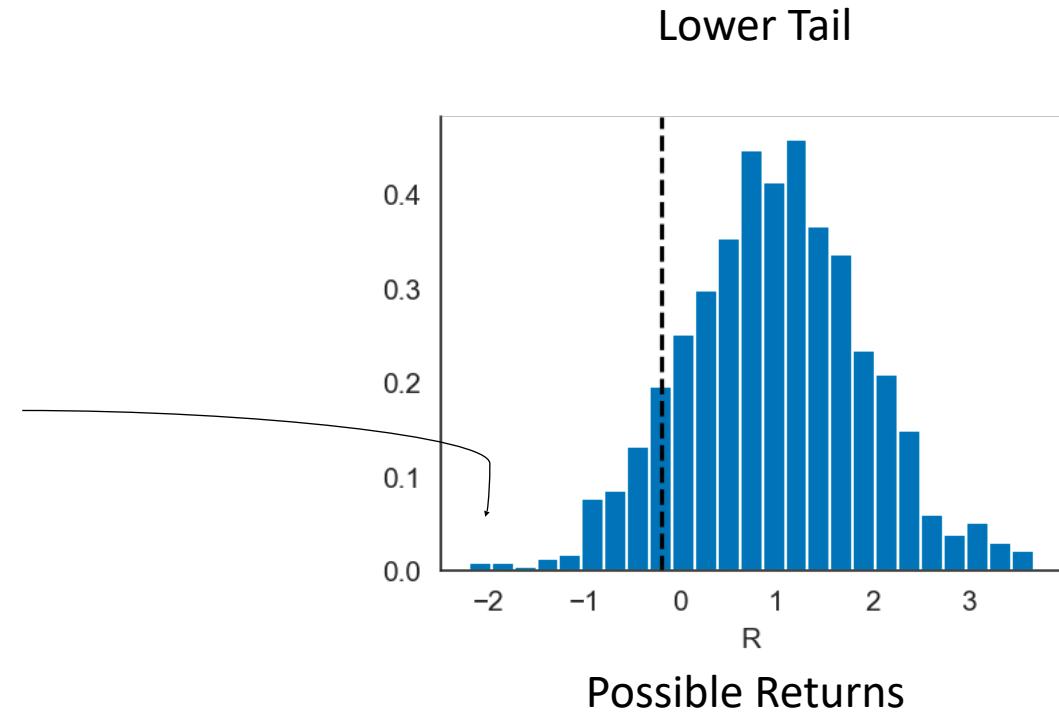
Decision making and risk

- risk is a critical aspect of decision making
- involves decision-making with respect to uncertain (probabilistic) outcomes
- industries have been designed around it (e.g. insurance markets)
- likely plays a crucial role in psychopathology (e.g. anxiety, mania) – ruminative ‘what-ifs’



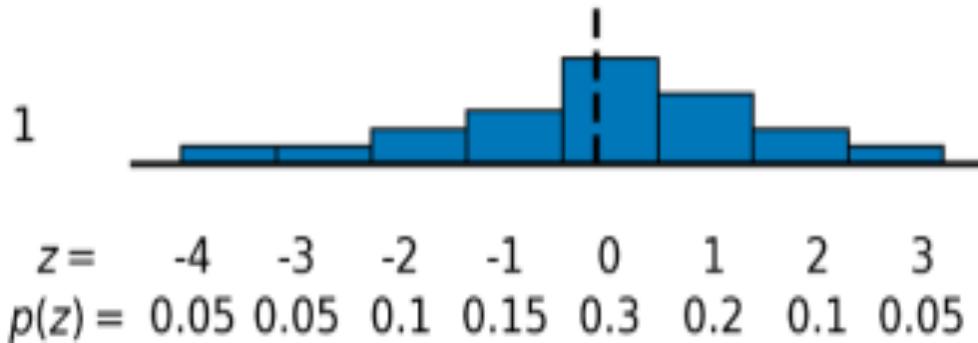
A useful risk measure from finance...

- caring about worst-case outcomes is natural in medicine, finance, engineering
- perhaps surviving predation for animals

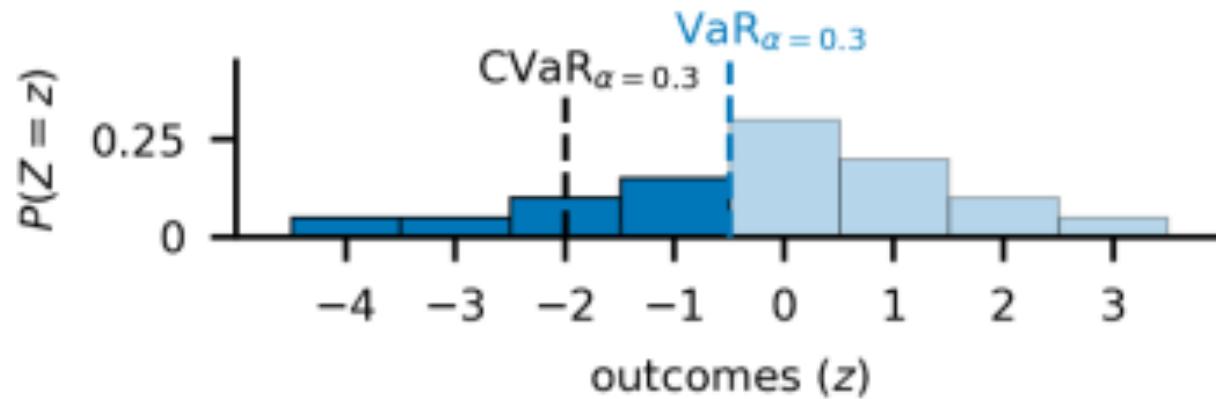


Modern Risk Measure: Conditional Value at Risk

- average case

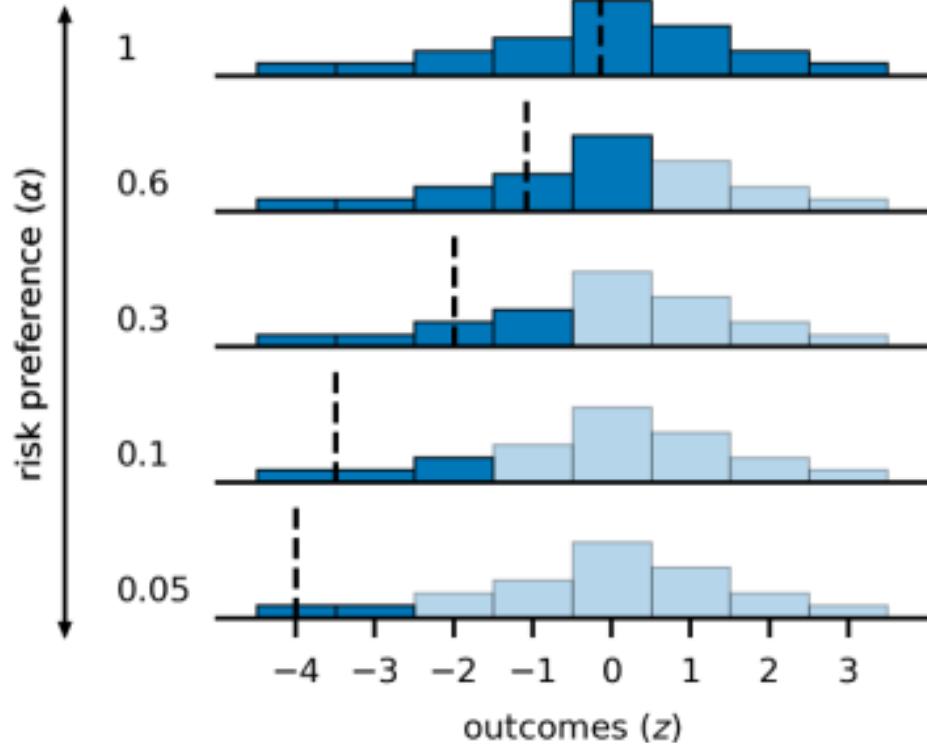


- worst α case: VaR
 - mean: CVaR



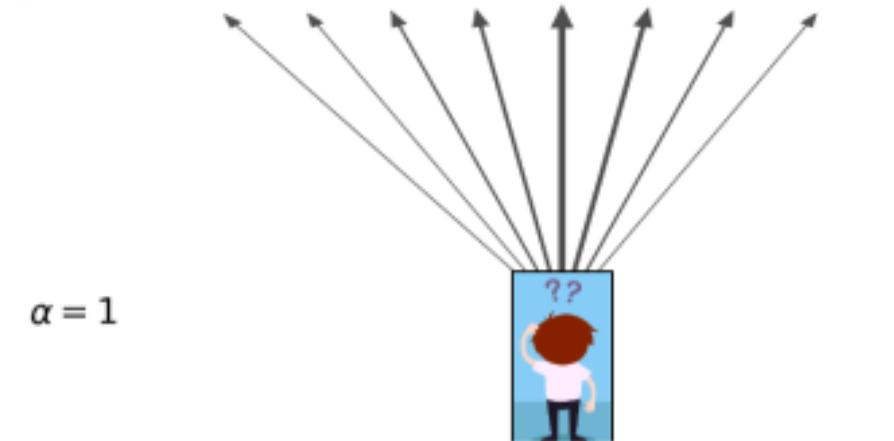
Two Views

(b)

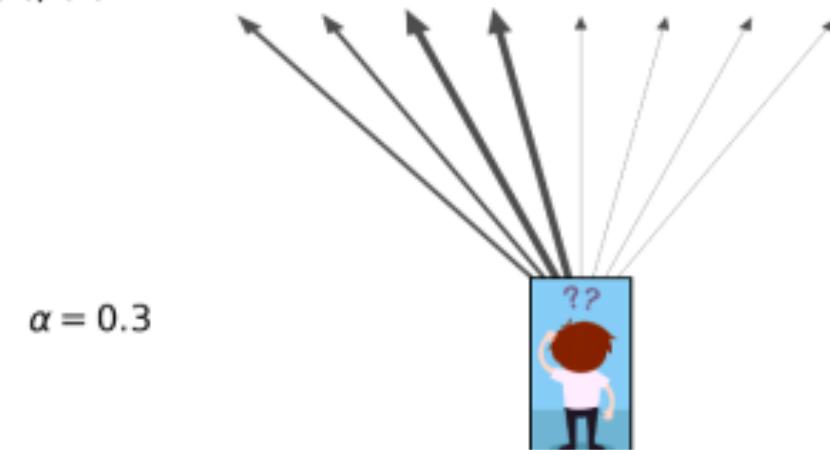


$$\text{CVaR}_\alpha[Z] = E[Z | Z \leq \text{VaR}_\alpha(Z)]$$

$$z = \begin{matrix} -4 & -3 & -2 & -1 & 0 & 1 & 2 & 3 \end{matrix}$$
$$p(z) = \begin{matrix} 0.05 & 0.05 & 0.1 & 0.15 & 0.3 & 0.2 & 0.1 & 0.05 \end{matrix}$$

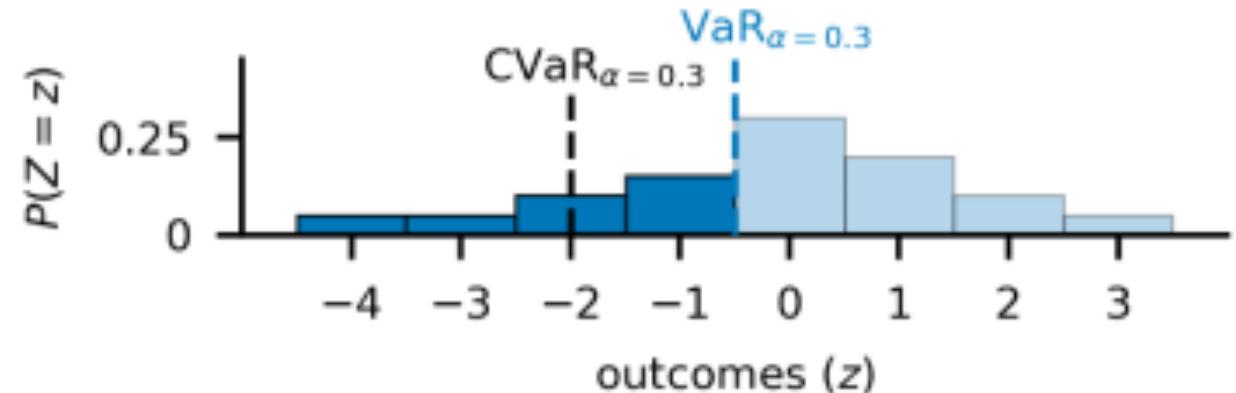


$$z = \begin{matrix} -4 & -3 & -2 & -1 & 0 & 1 & 2 & 3 \end{matrix}$$
$$\xi(z)p(z) = \begin{matrix} 0.16 & 0.16 & 0.33 & 0.33 & 0.0 & 0.0 & 0.0 & 0.0 \end{matrix}$$



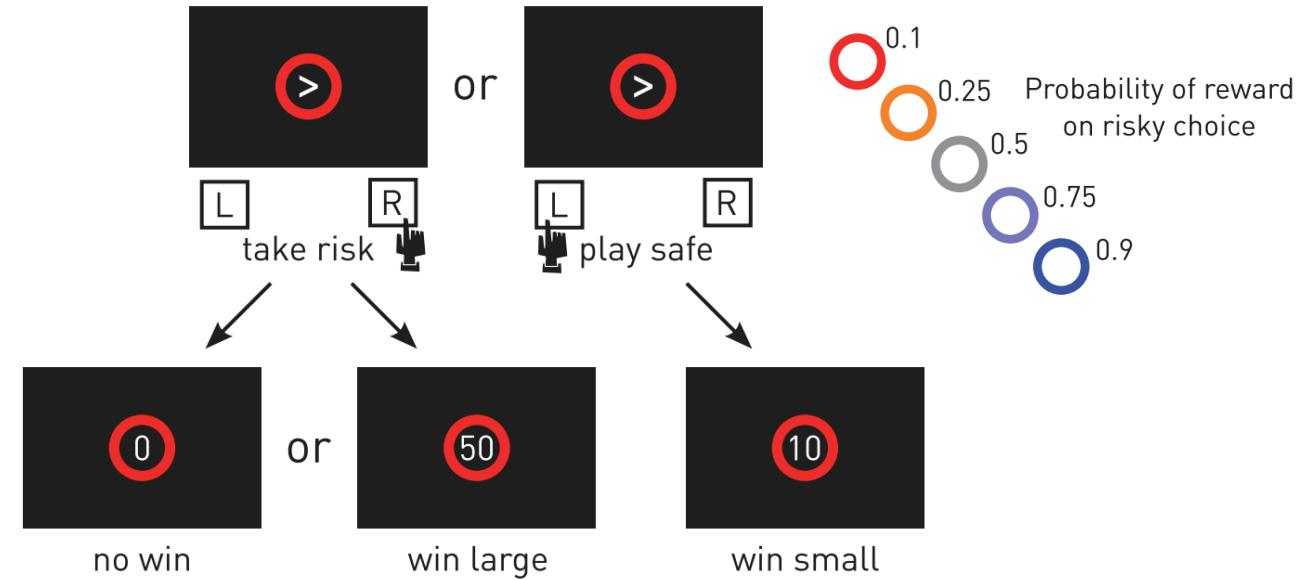
CVaR: Conditional Value at Risk

- coherent risk measure
- emphasizes the lower tail
- $\alpha = 1$: the ‘regular’ mean
- $\alpha \searrow 0$: worst case – the minimum
- equivalent to distorted probabilities favouring bad outcomes

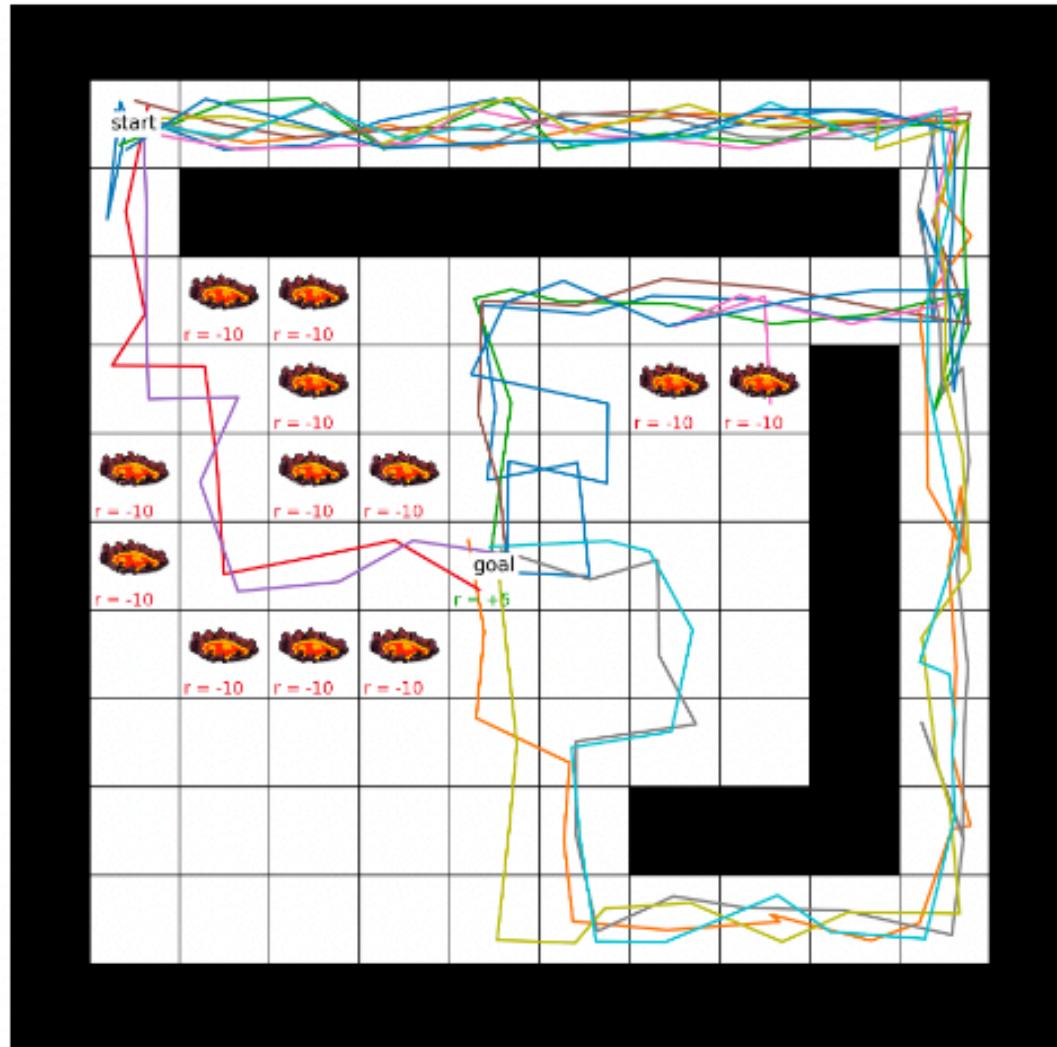


Experimental paradigms

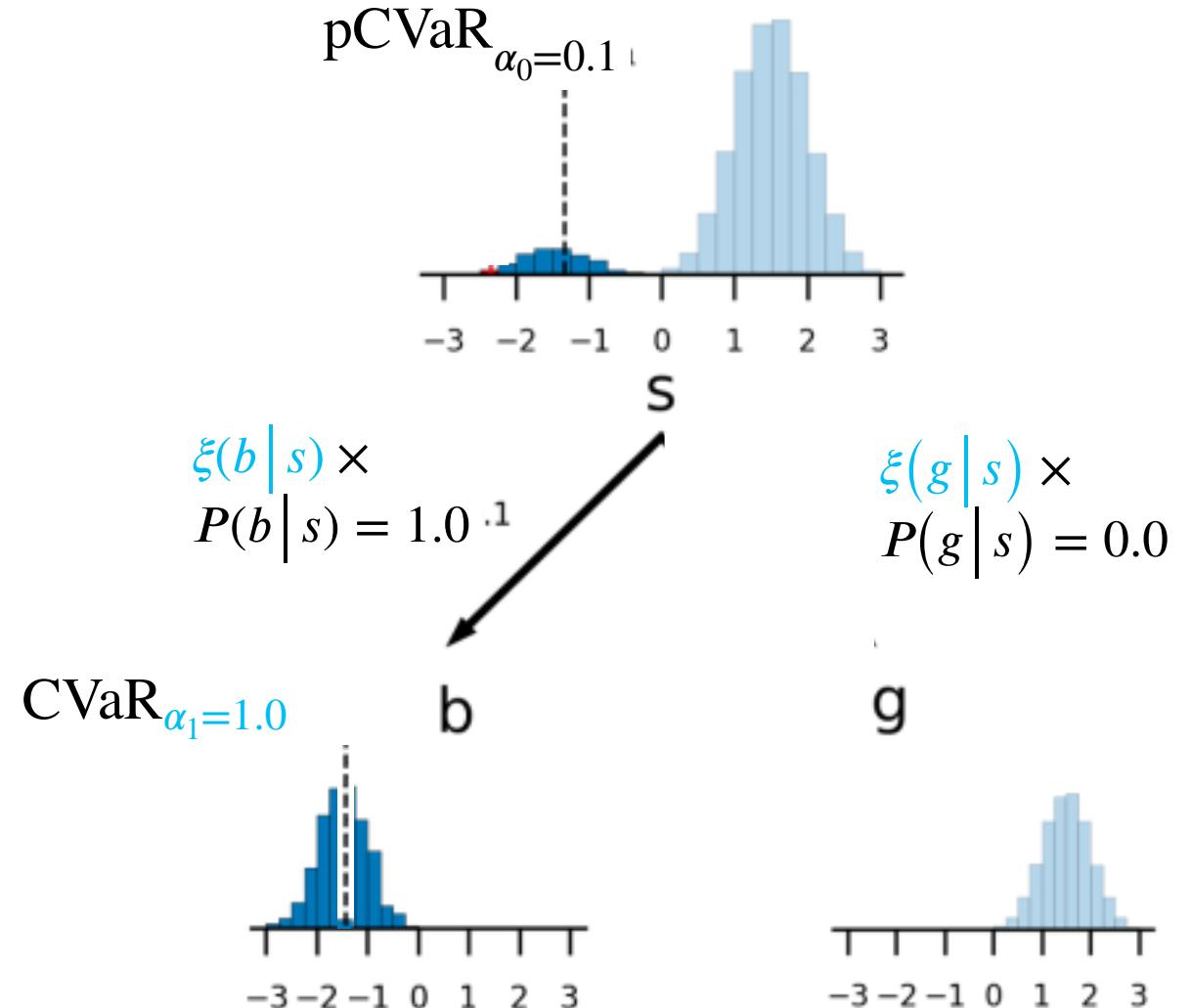
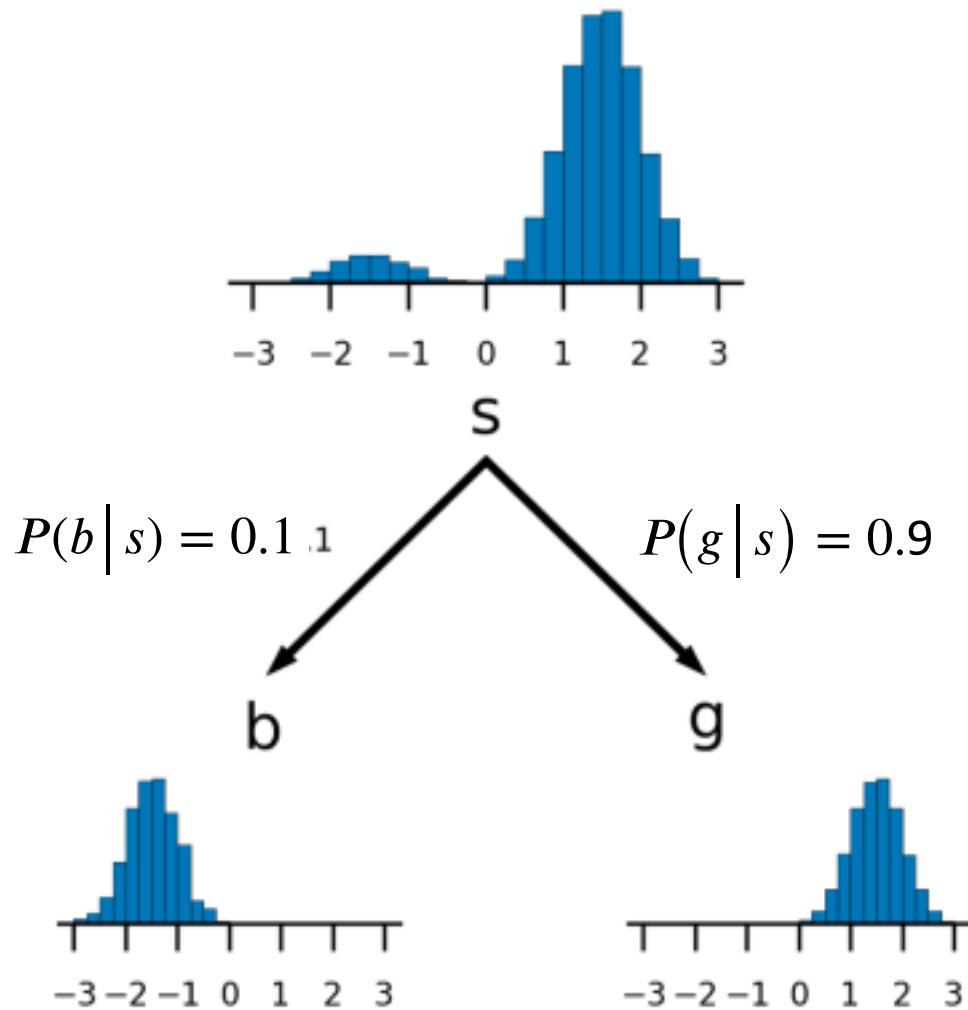
- usually:



What about sequential choice?



What about the sequential case? V1: pCVaR

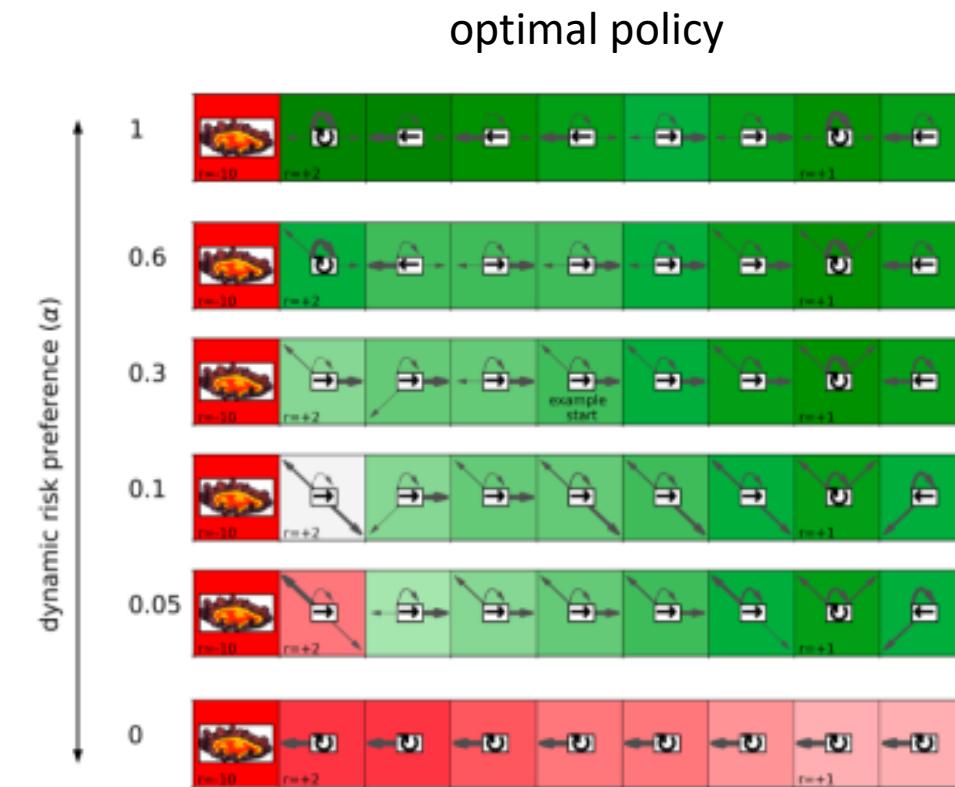
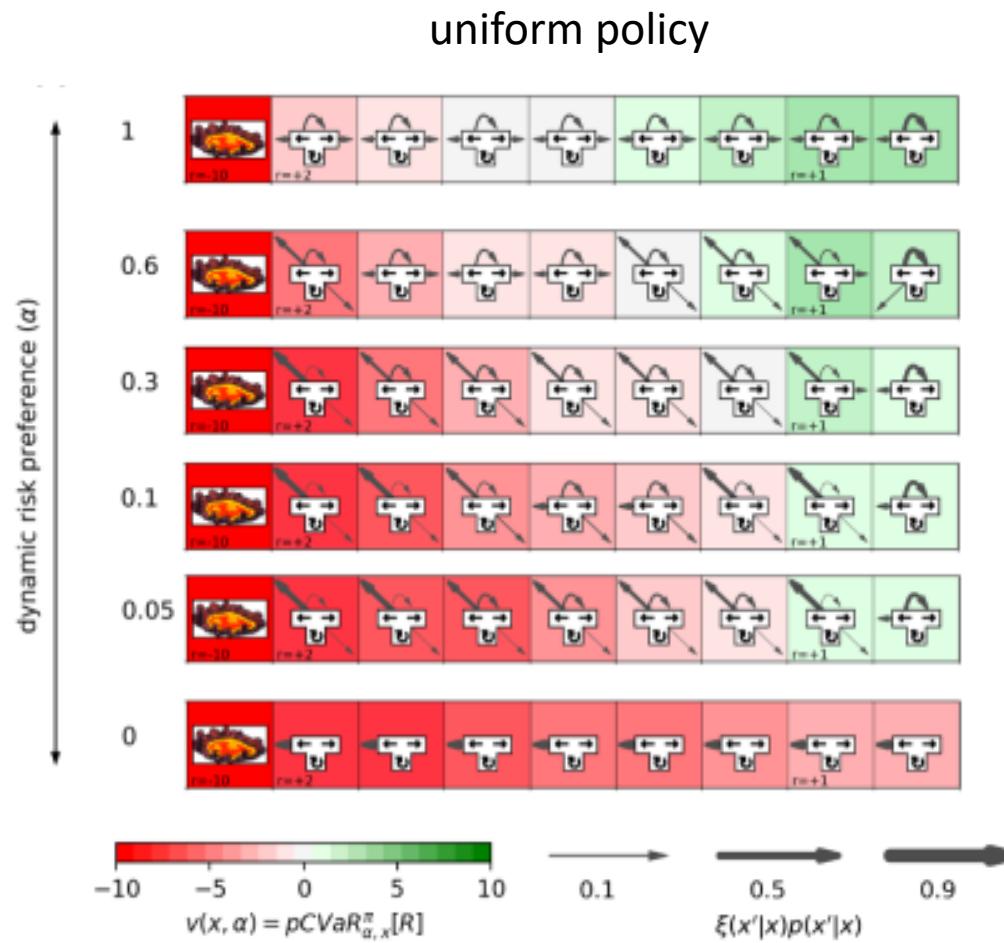
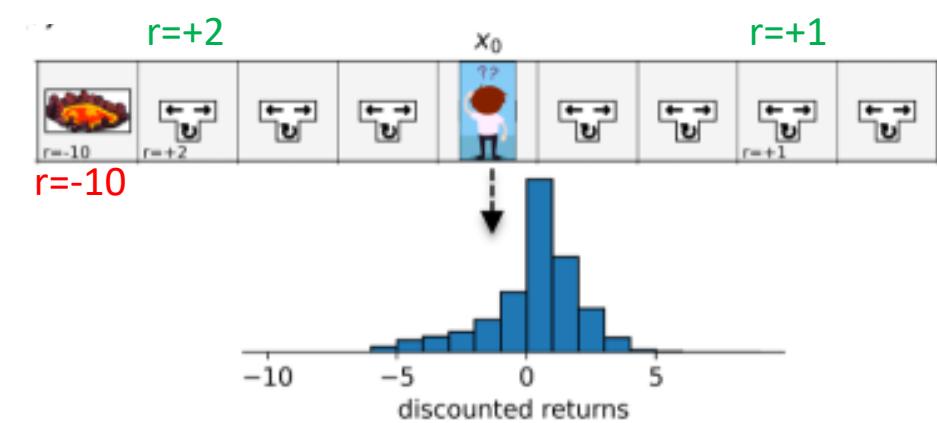


Sequential pCVaR

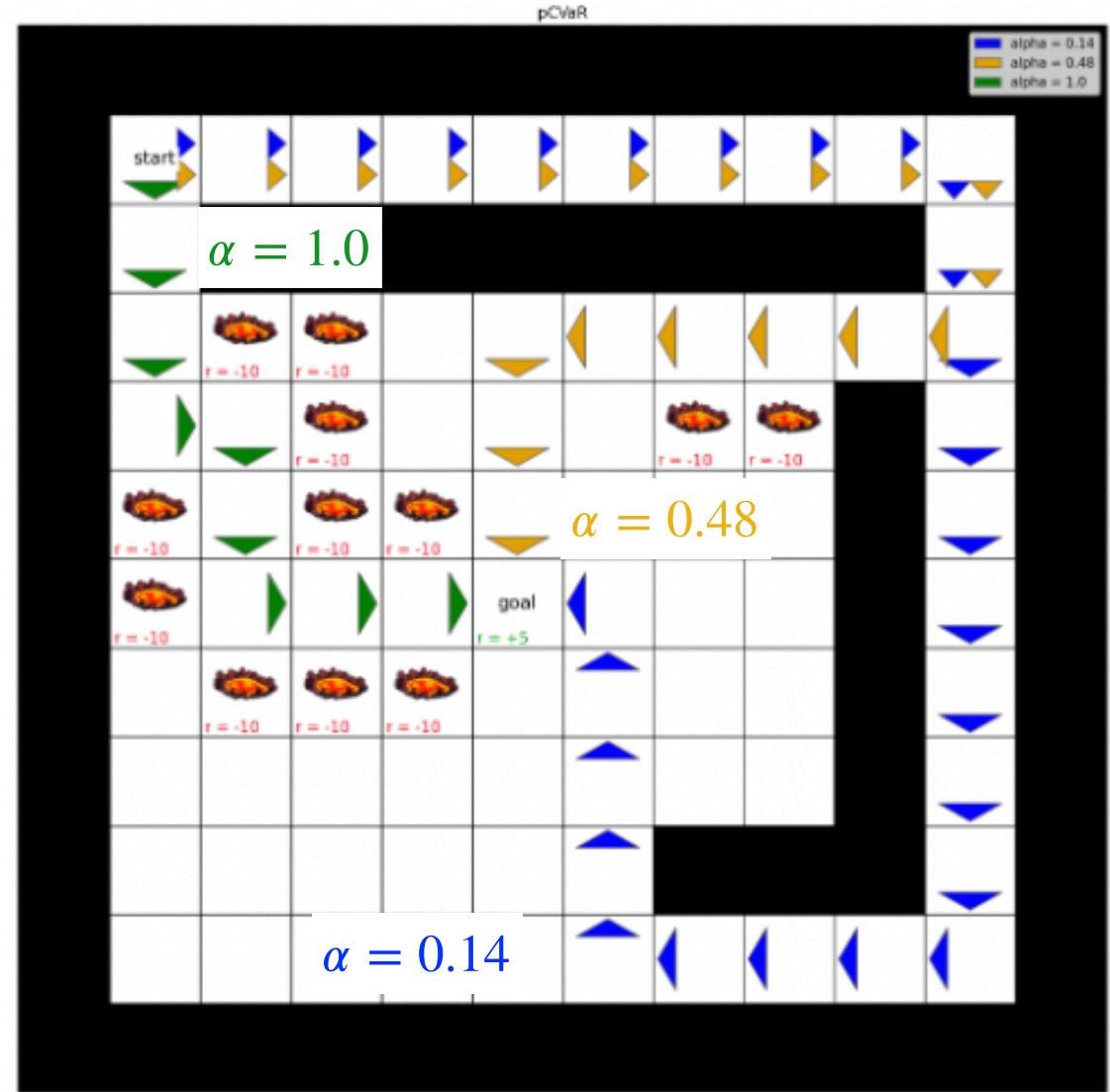
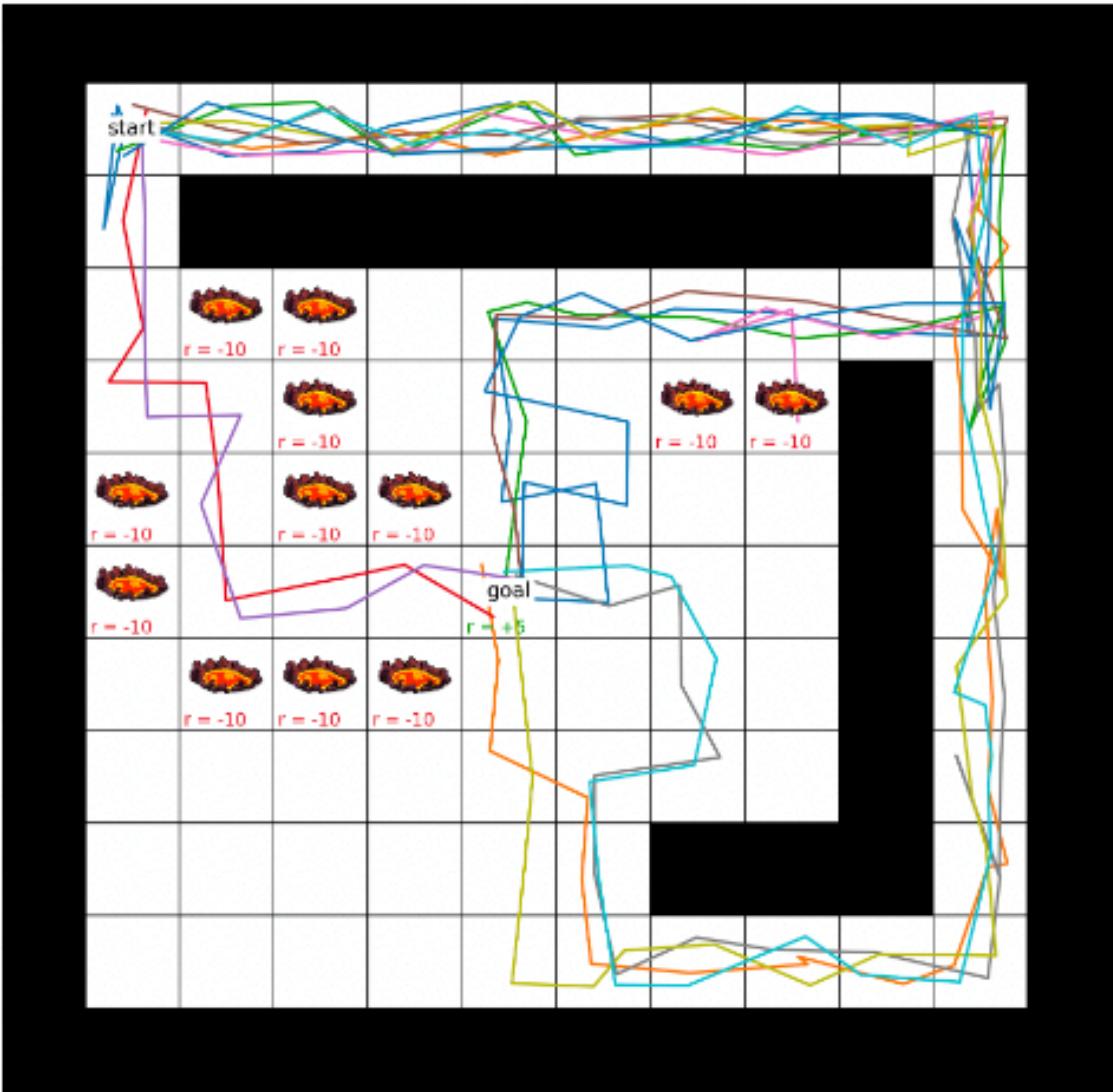
$$\text{pCVaR}_{\alpha_0, x_0}^{\pi}[R] := \text{CVaR}_{\alpha_0}[R_0 + \gamma R_1 + \gamma^2 R_2 \dots | X_0 = x_0, \pi]$$

- precommitted CVaR: pCVaR
 - privilege a start state: home; nest
 - change α according to the gambler's fallacy
 - if unlucky: α increases $\alpha = 0; 1$ are special
 - if lucky: α decreases
 - either history-dependent evaluation
 - or add the α dimension with transitions tied to ξ

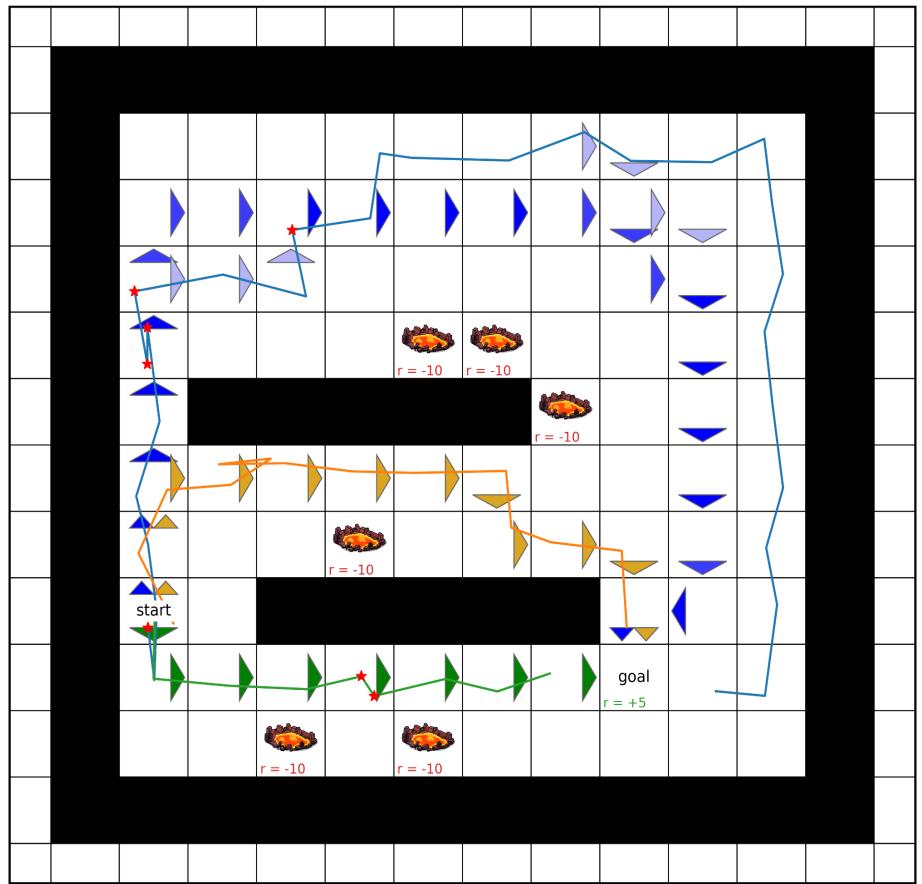
pCVaR in a random walk



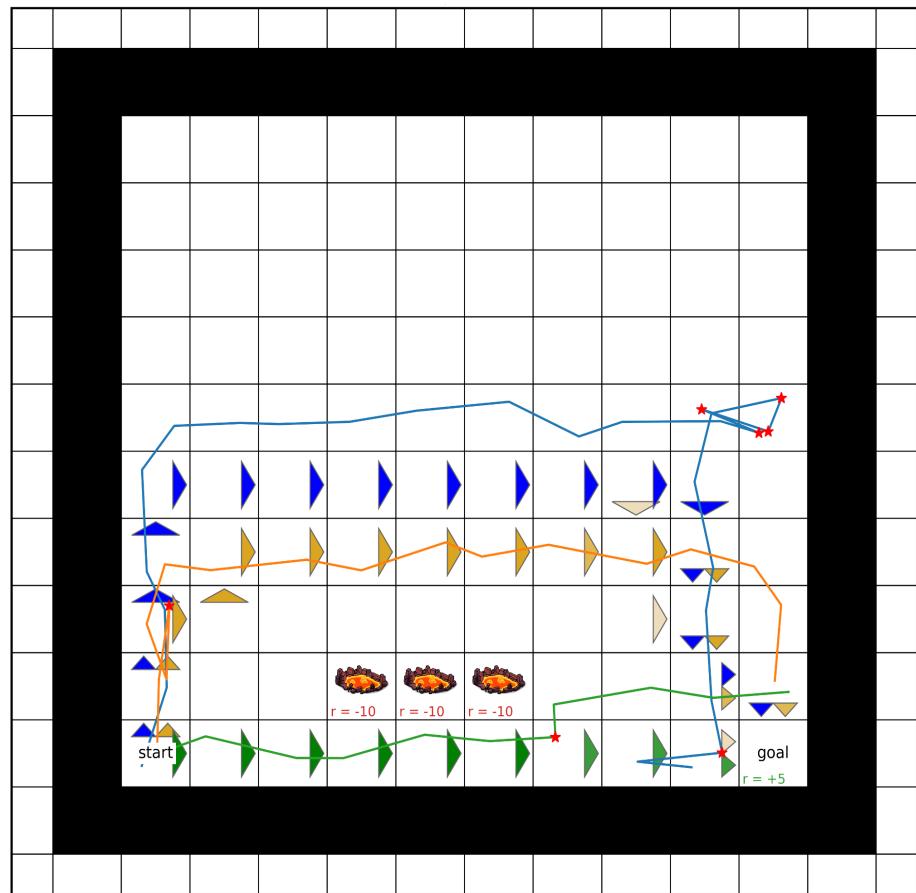
With the Lava Pits...



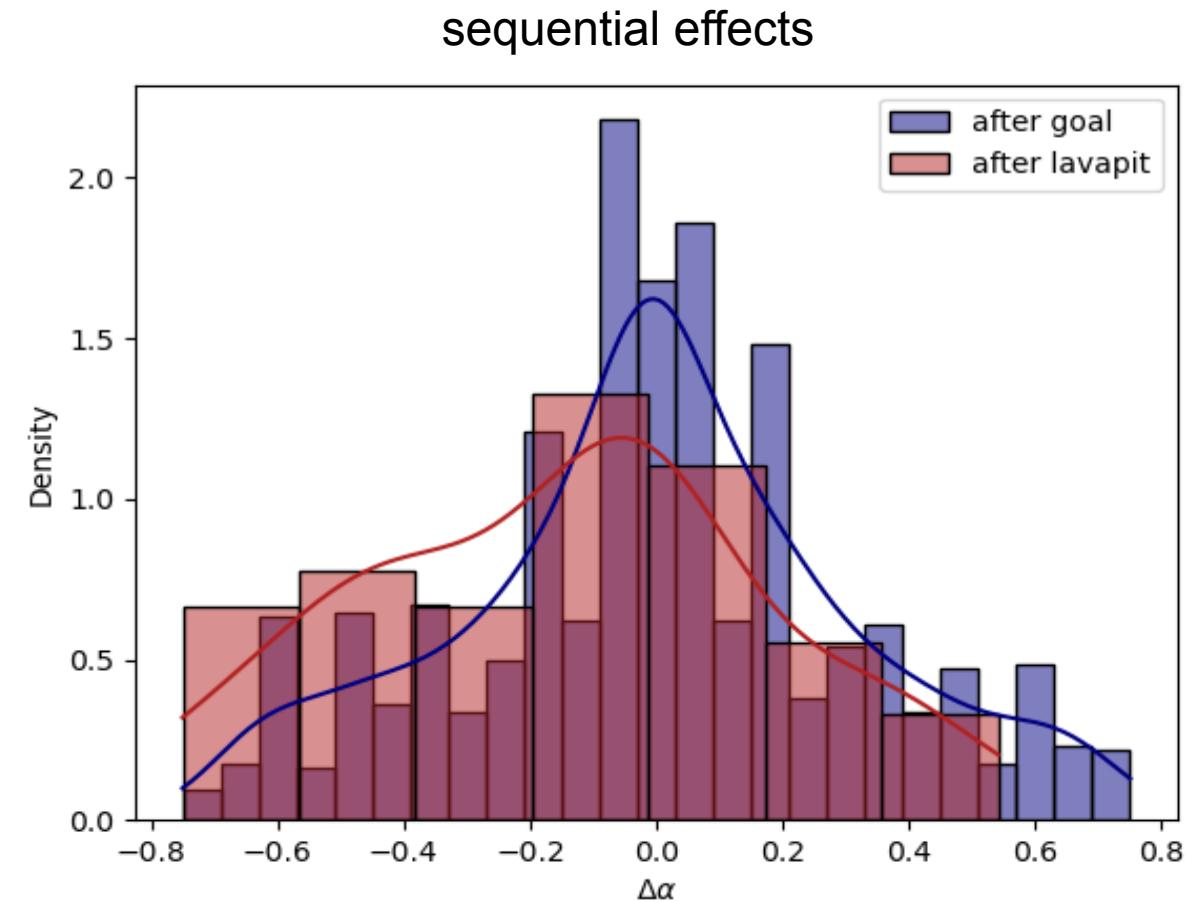
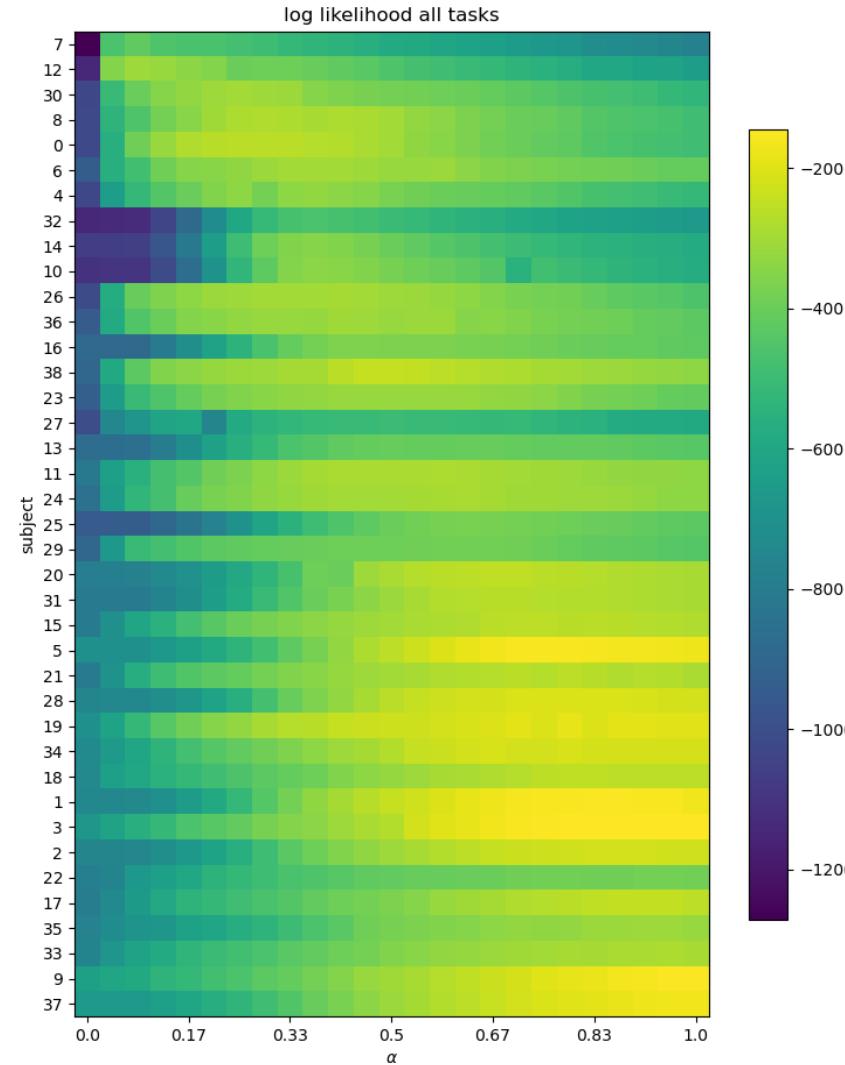
Consistent within a subject



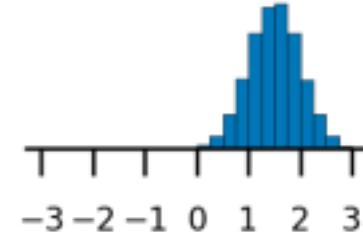
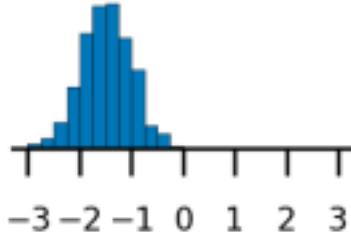
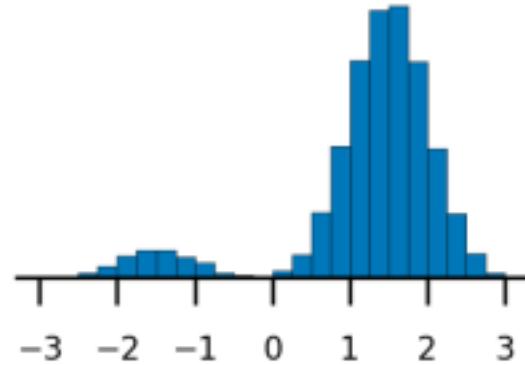
$\alpha = 0.21$
 $\alpha = 0.46$
 $\alpha = 0.88$



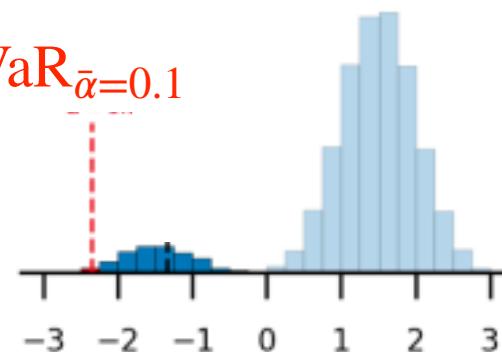
Risk-sensitivity across subjects



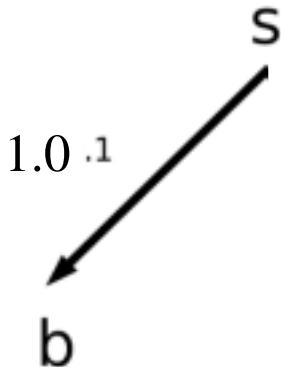
What about the sequential case? V2: nCVaR



nCVaR $_{\bar{\alpha}=0.1}$



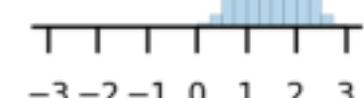
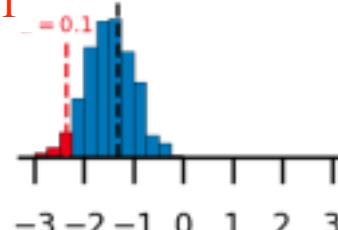
$$\xi(b|s) \times P(b|s) = 1.0 \cdot 0.1$$



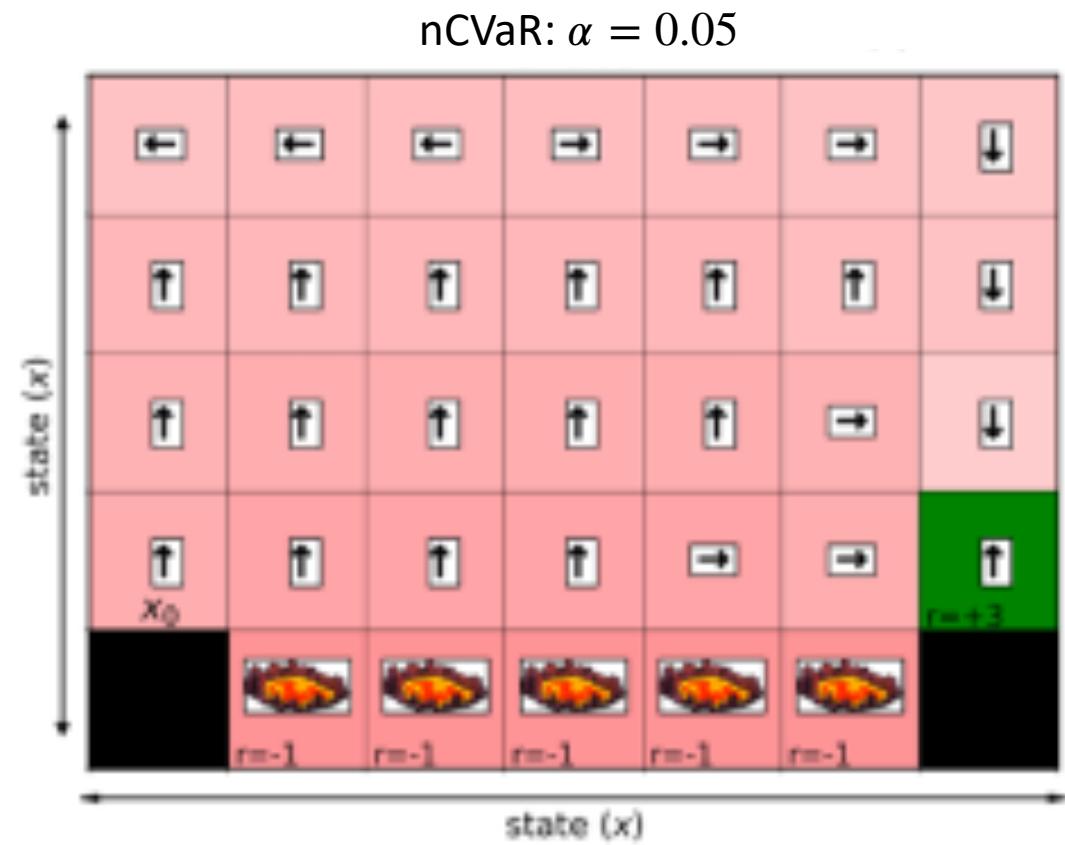
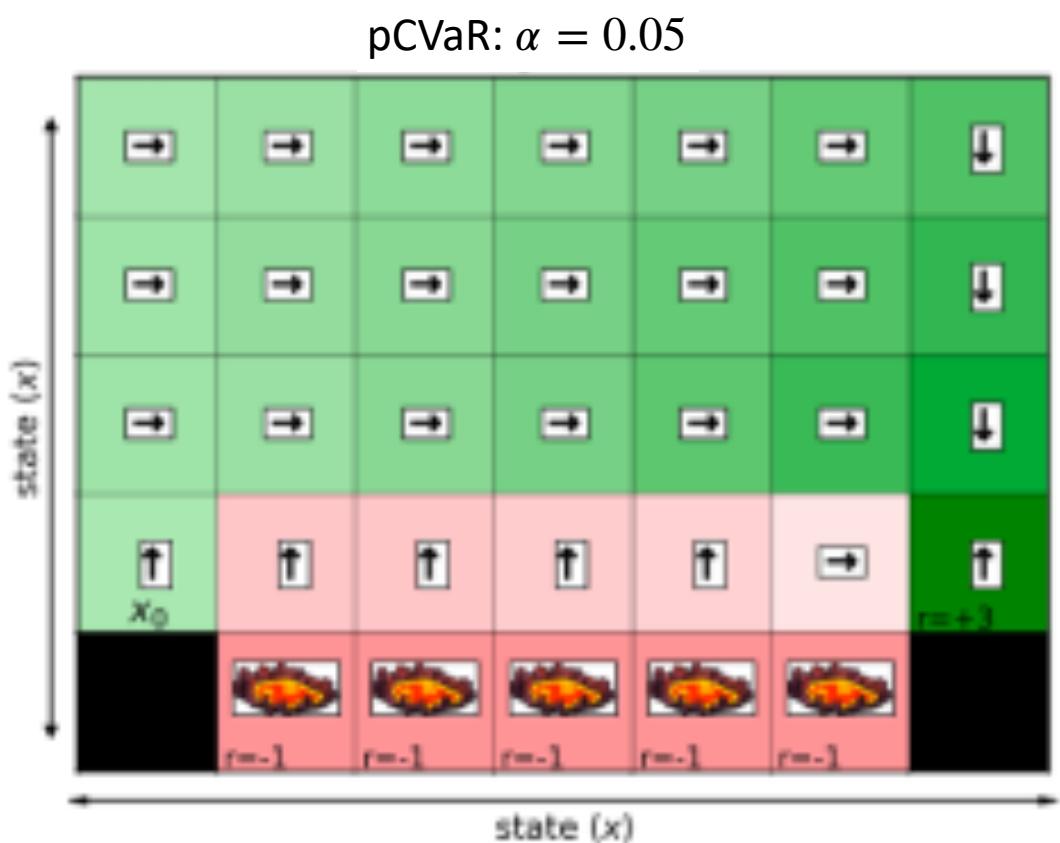
$$\xi(g|s) \times P(g|s) = 0.0$$

g

CVaR $_{\bar{\alpha}=0.1}$



nCVaR for the cliff



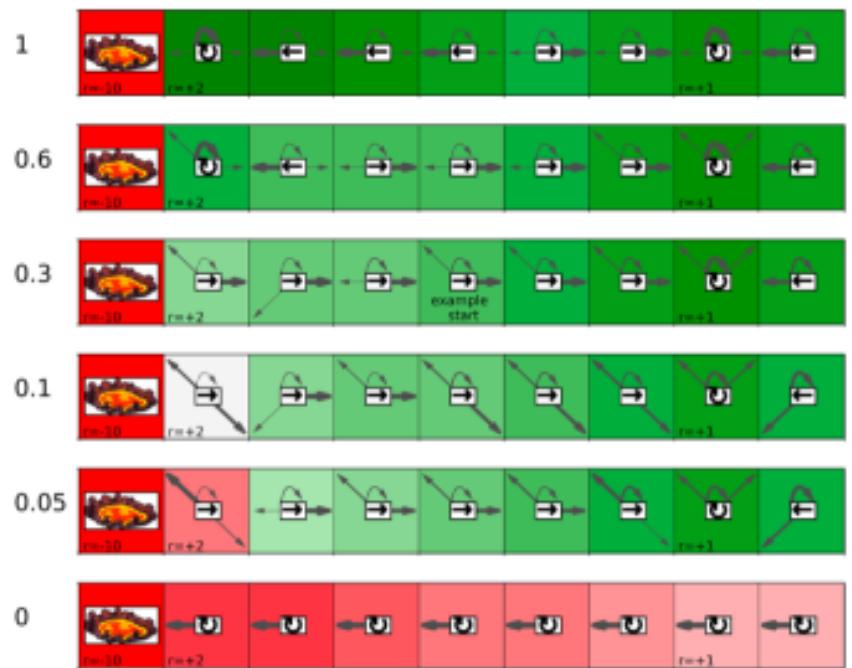
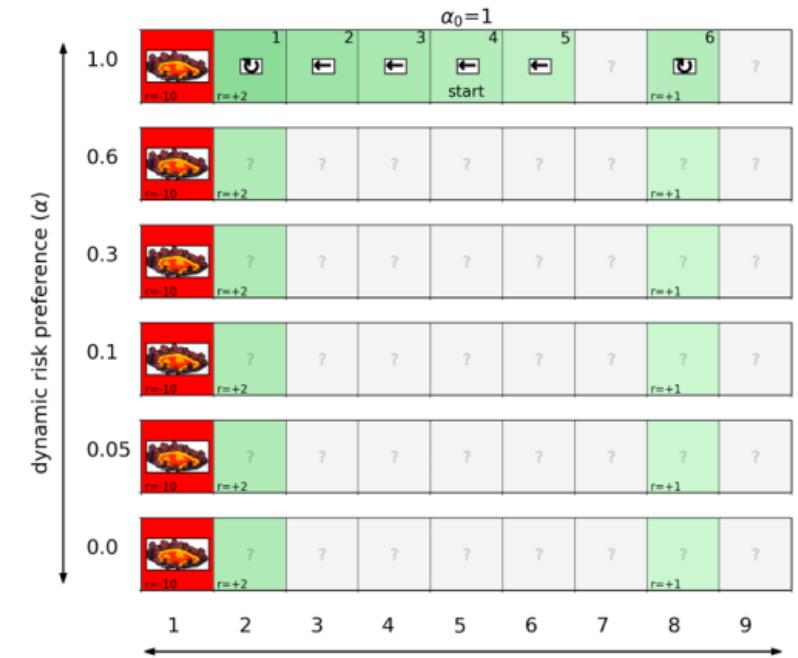
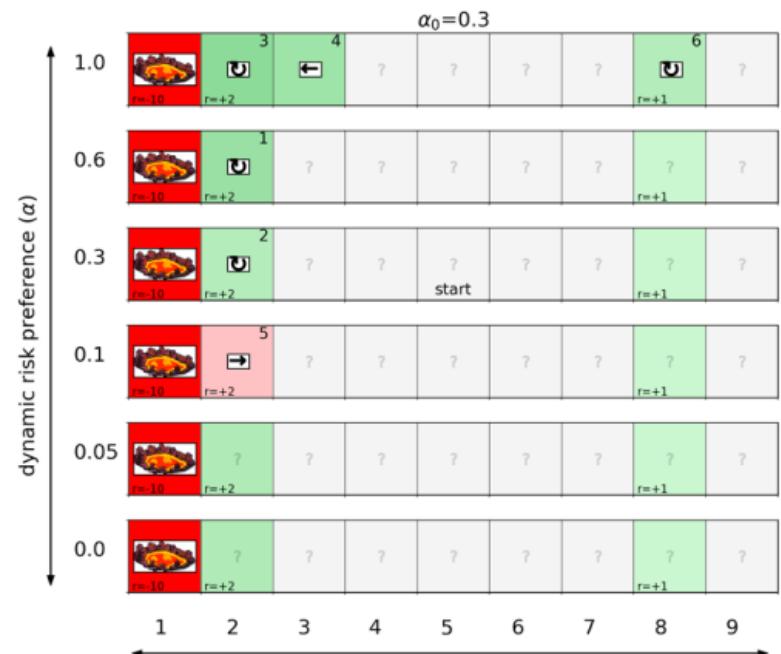
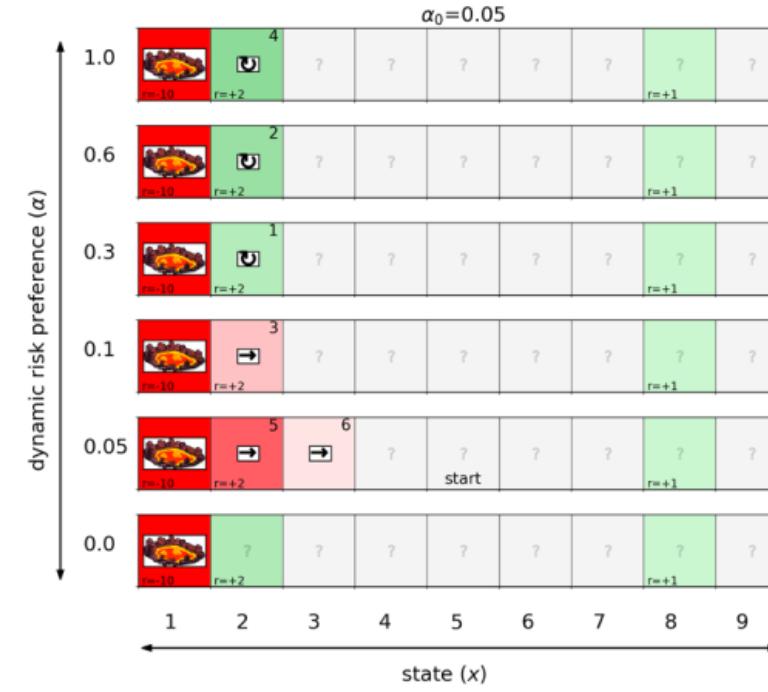
Interim summary

- parametric risk-avoidant behavior
 - from pre-committed pCVaR: with gambler's fallacy
 - more complex inference
 - other forms of risk avoidance (mean variance) also use extra dimension (current return)
 - from nested nCVaR: with excessive risk aversion
 - still need extra dimension to adjust risk aversion
 - psychiatrically
 - 'wrong problem': pathological avoidance
 - stay at home
 - nCVaR makes this worse
 - in stochastic problems, $\alpha = 0$ leads to indifference/helplessness

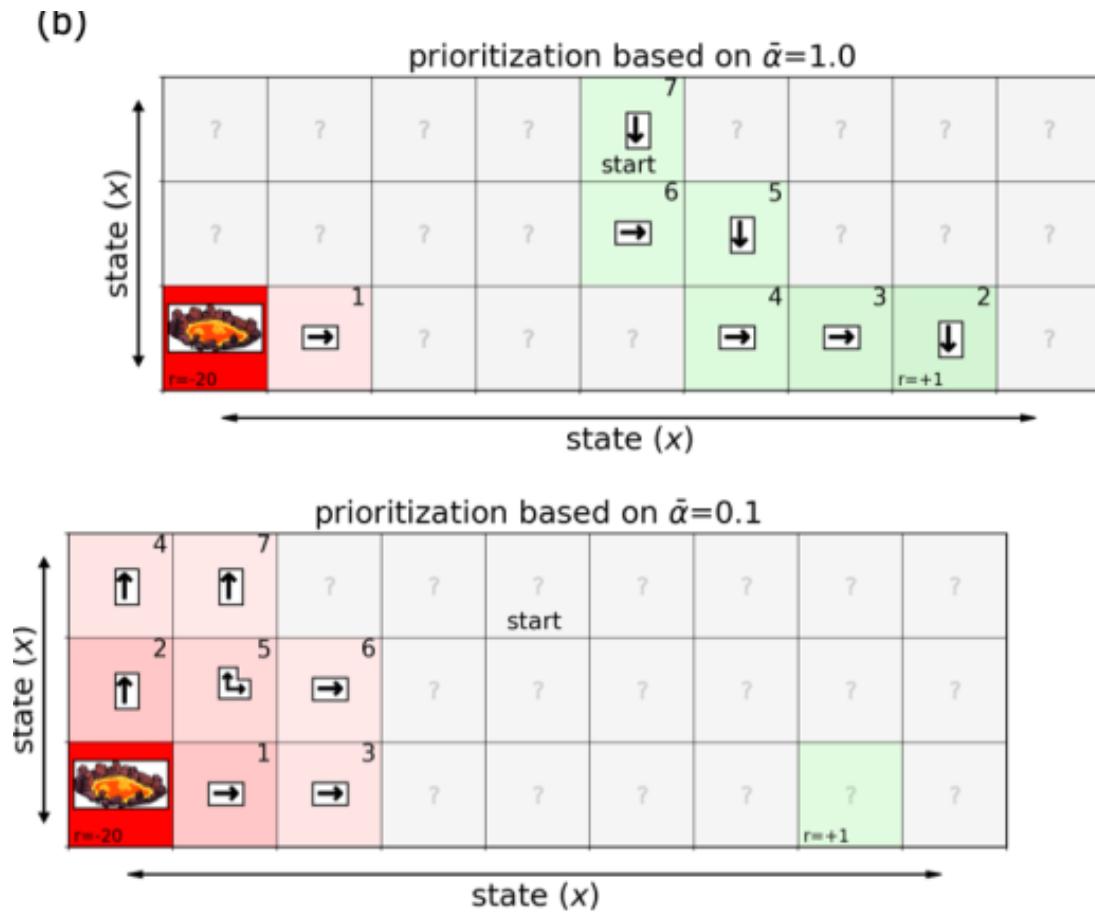
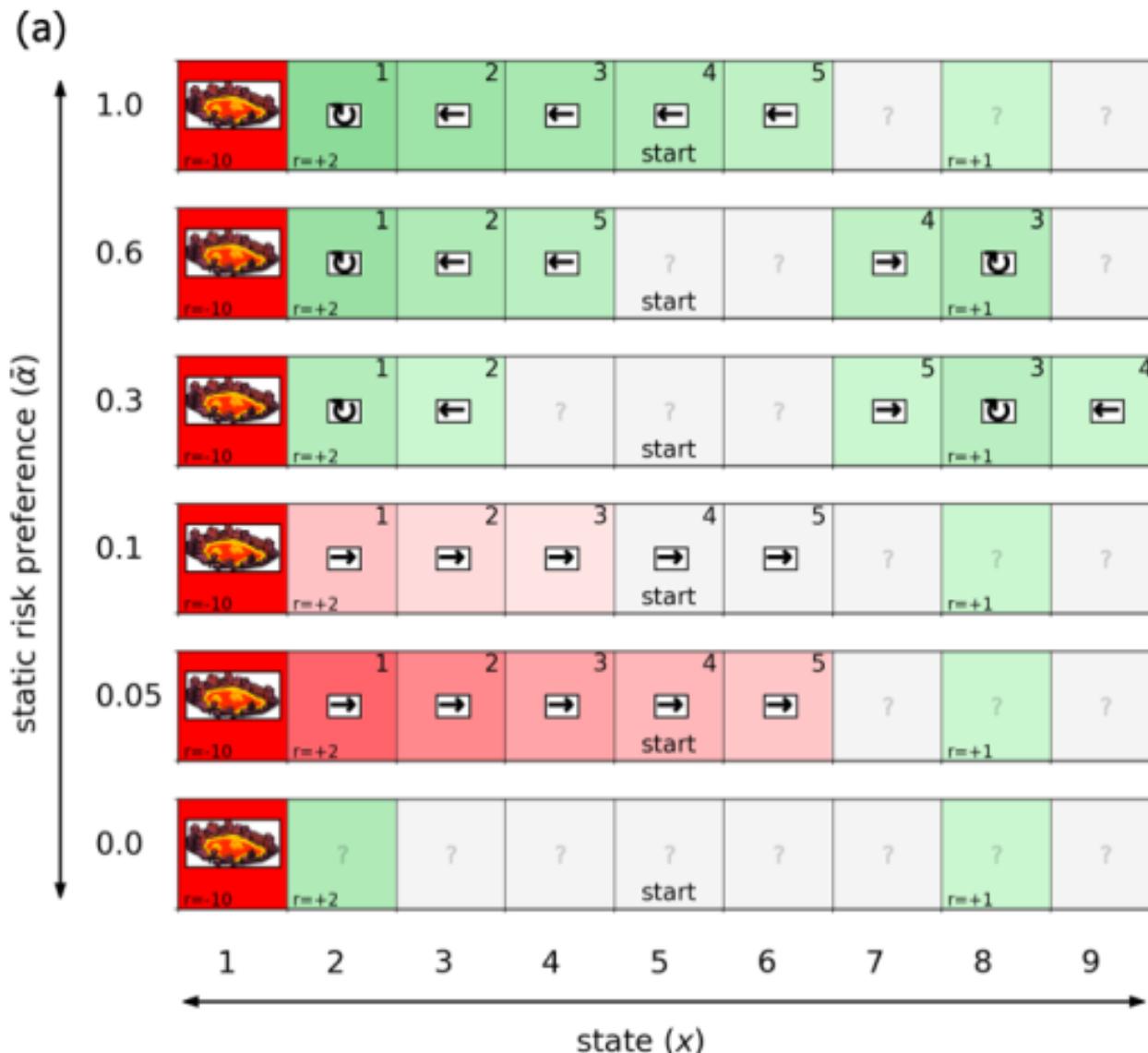
Risk-avoidant planning

- planning:
 - on-line: model-based reinforcement learning (Monte-Carlo tree search)
 - off-line:
 - use coordinated hippocampal/cortical replay to invert the generative task model
 - evidence in rodents and humans
 - RL:
 - DYNA: Sutton – to enable exploration
 - Mattar & Daw – synergize with prioritized sweeping (Moore)
 - choose to update based on: product of
 - gain: how much you change your policy based on an update
 - need: how frequently you will visit that state in the future
 - unreasonably: assume optimal calculations in value iteration

Optimal planning for pCVaR



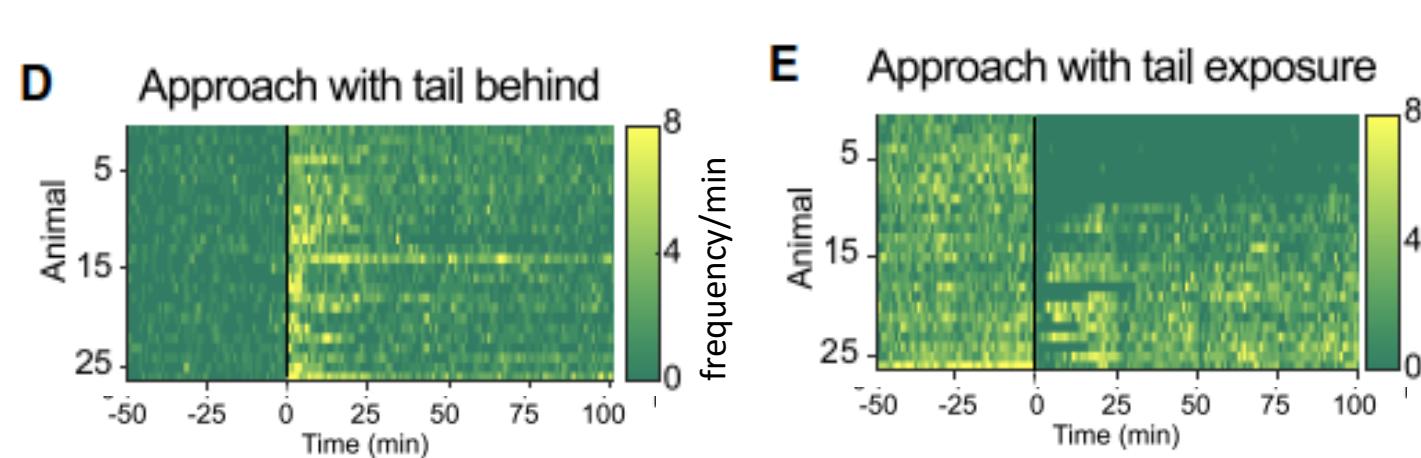
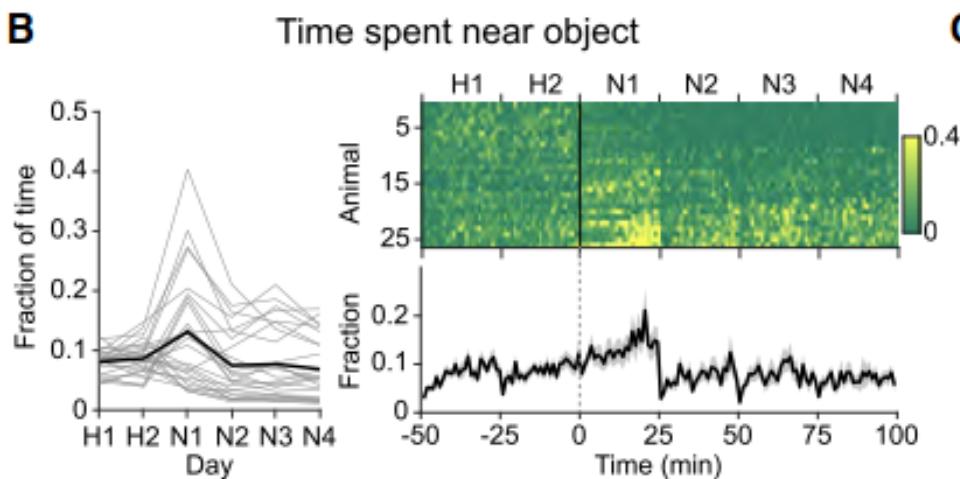
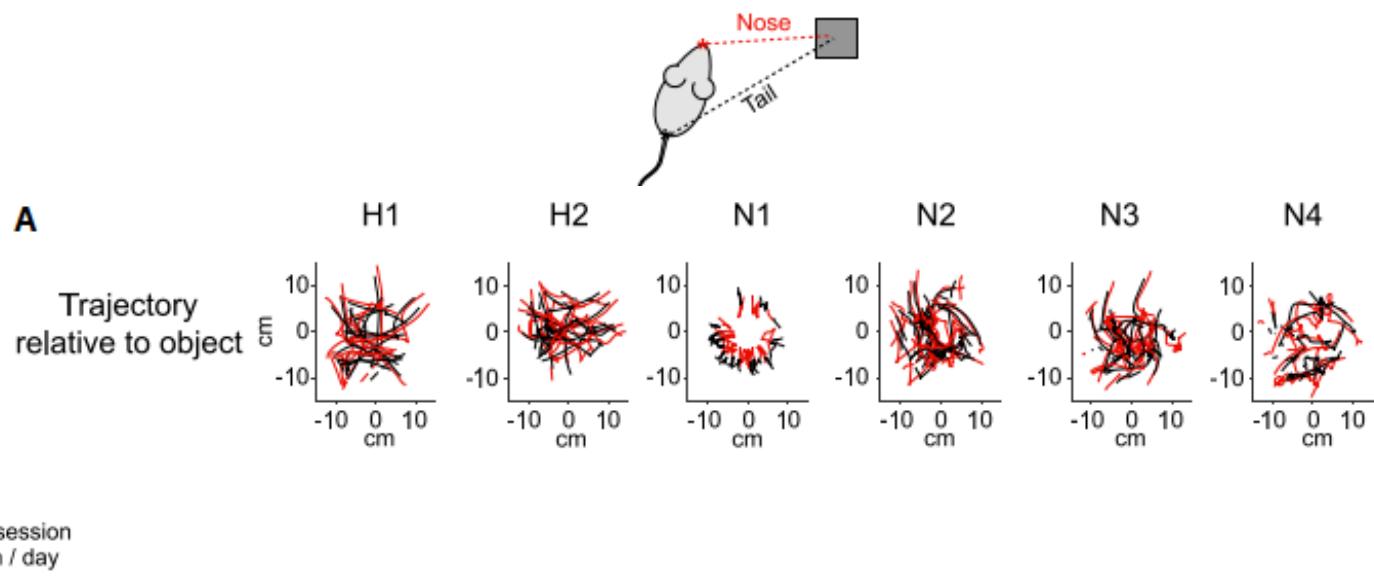
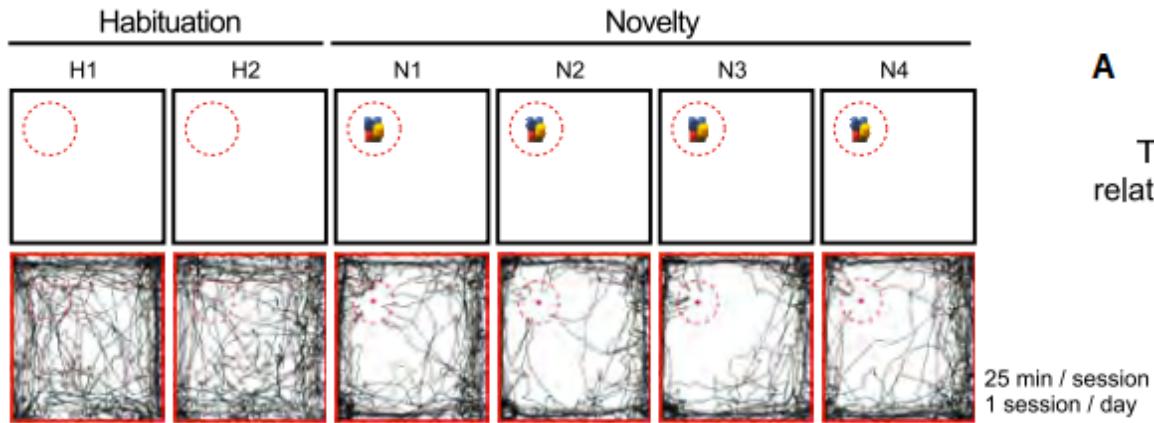
Optimal planning for nCVaR



Discussion

- wrong problem:
 - *optimally* dysfunctional avoidance and rumination from low α and nCVaR
 - action indifference and helplessness from α near 0
 - threshold for improvement in CVaR from planning: meta-control
 - non-parametric Bayesian (infinite) model – always another catastrophe around the corner
 - effects on exploration/exploitation trade-off
 - robustness to mis-specification
- wrong solution:
 - ineffective updates – for instance from Pavlovian avoidance
 - serotonergic pruning?
 - pCVaR with incomplete adjustment for ‘luck’ – continuity to nCVaR
- wrong environment:
 - over-generalizing representations

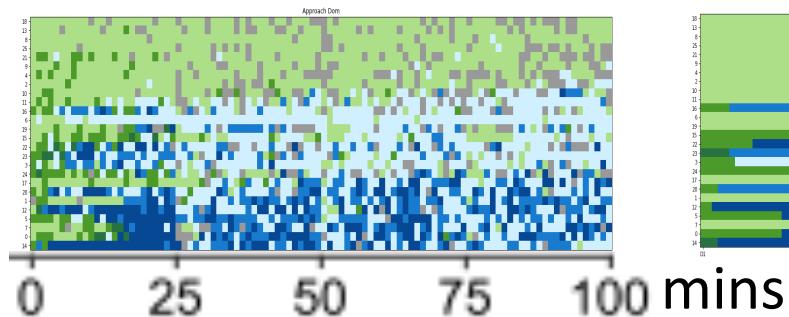
Not only humans...



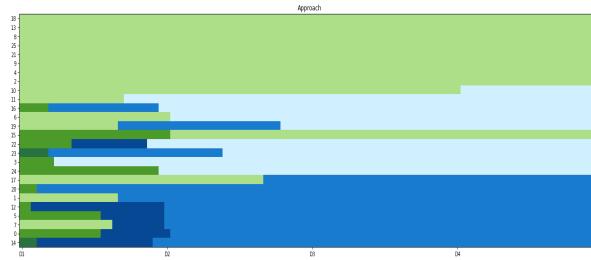
Animal Data

frac time <7cm

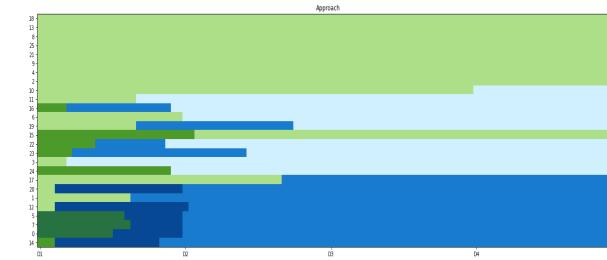
Minute-to-minute



Phase-wise



Model Fit



Time (Cau)
minute/minute Time
(Con)

0.2+	" "
0.1 - 0.2	" "
1e-6 - 0.1	" "