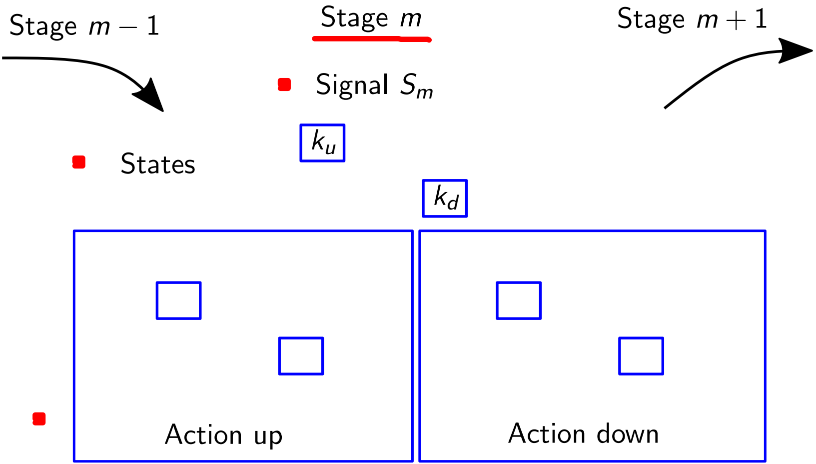
Easy strategies in complex games

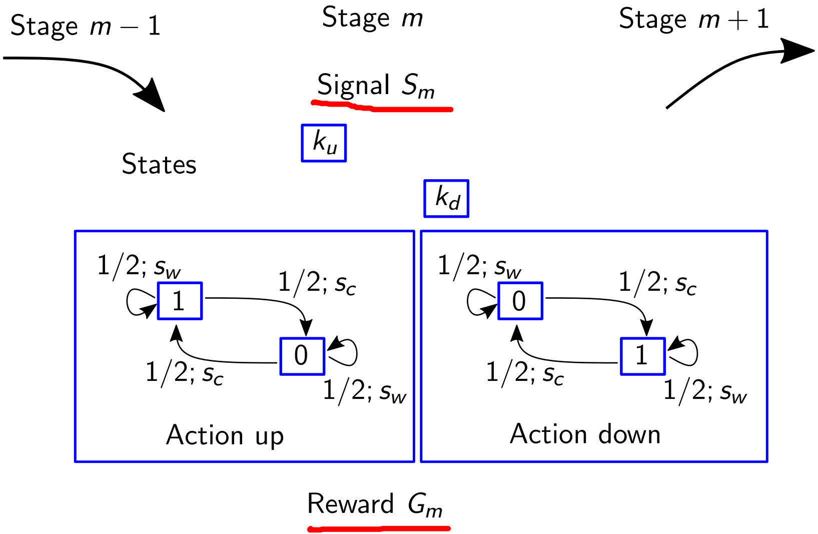
Finite memory strategies in POMDPs with long-run average objective

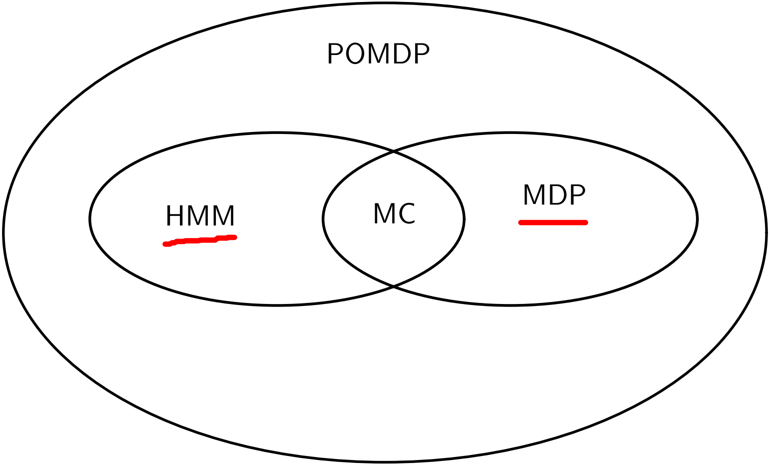
K. Chatterjee[1] **R. Saona**[1] B. Ziliotto[2]

[1]IST Austria

[2]CEREMADE, CNRS, Université Paris Dauphine, PSL Research Institute


Stage $m-1$
Stage m
Stage $m+1$
Signal S_m
k_u
States
k_d
Action up
Action down


Stage $m-1$
Stage m
Stage $m+1$
Signal S_m
k_u
States
k_d
1/2; s_w
1
1/2; s_c
0
1/2; s_c
1/2; s_w
Action up
1/2; s_w
0
1/2; s_c
1
1/2; s_c
1/2; s_w
Action down
Reward G_m


POMDP
HMM
MC
MDP

Stage $m-1$

Stage m

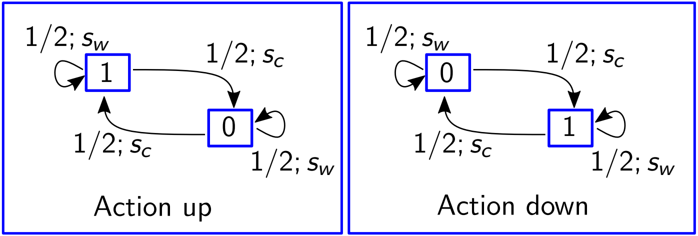
Stage $m+1$

Signal S_m

States

$1/4$ k_u

$3/4$ k_d



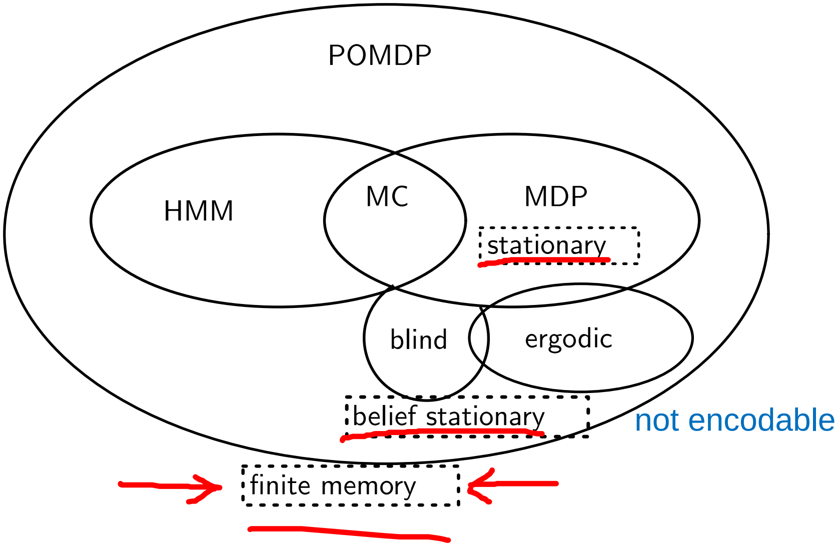
Reward G_m

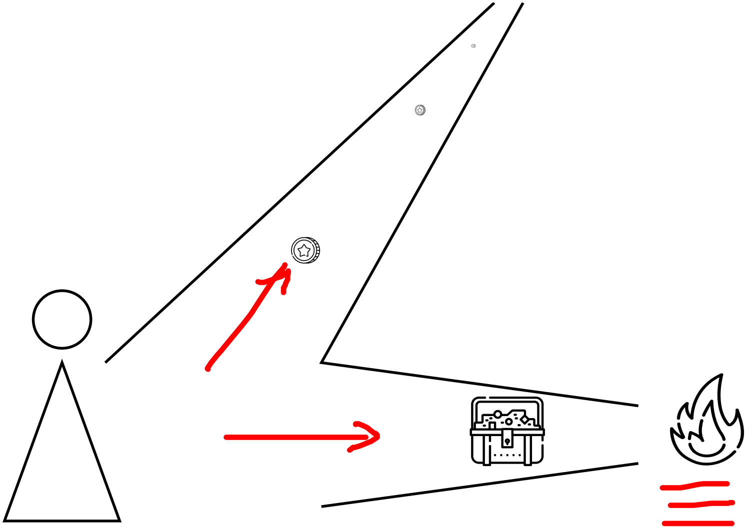
$$v_\infty(p_1) := \sup_{\sigma \in \Sigma} \mathbb{E}_\sigma^{p_1} \left(\liminf_{n\to\infty} \frac{1}{n} \sum_{m=1}^{n} G_m \right)$$

$$= \lim_{n\to\infty} \sup_{\sigma \in \Sigma} \mathbb{E}_\sigma^{p_1} \left(\frac{1}{n} \sum_{m=1}^{n} G_m \right)$$

$$= \lim_{\lambda\to 0^+} \sup_{\sigma \in \Sigma} \mathbb{E}_\sigma^{p_1} \left(\sum_{m=1}^{\infty} \lambda(1-\lambda)^{m-1} G_m \right)$$

$$= \liminf_{n\to\infty} \sup_{\sigma \in \Sigma} \mathbb{E}_\sigma^{p_1} \left(\frac{1}{n} \sum_{m=1}^{n} G_m \right)$$


POMDP
HMM
MC
MDP
stationary
blind
ergodic
belief stationary
not encodable
finite memory



Icons made by Freepik from www.flaticon.com

Approximation.

$$|v - v_\infty(p_1)| \leq \varepsilon \,.$$

undecidable

This is impossible.

Lower bound.

$$(v_n) \nearrow v_\infty(p_1).$$

Our result.

Upper bound.

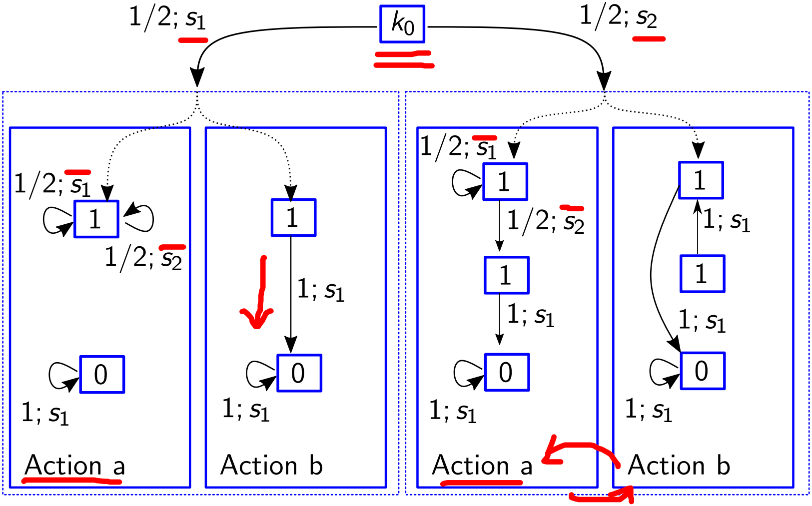
$$(v_n) \searrow v_\infty(p_1).$$

This is impossible.

Continuity(?).

$$v_\infty(p_1) = F(\text{rewards}, \text{transitions})\,.$$

Continuous with respect to rewards and lower semi-continuous with respect to transitions.



Property. We need to recall the first signal to play ε-optimally.

Continuity.

$$v_\infty(p_1) = F(\text{rewards}, \text{transitions}) \,.$$

Is v_∞ continuous with respect to transitions?

Belief partition.

$$v_\infty(p_1) = \sup_{\sigma \in ??} \mathbb{E}_\sigma^{p_1} \left(\liminf_{n \to \infty} \frac{1}{n} \sum_{m=1}^{n} G_m \right) \,.$$

Do belief partition strategies have this property?