Applicable Adaptive Discounted Fully Probabilistic Design of Decision Strategy

Soňa Molnárová

Faculty of Nuclear Sciences and Physical Engineering
Czech Technical University

June 21st 2024

General Problem Description

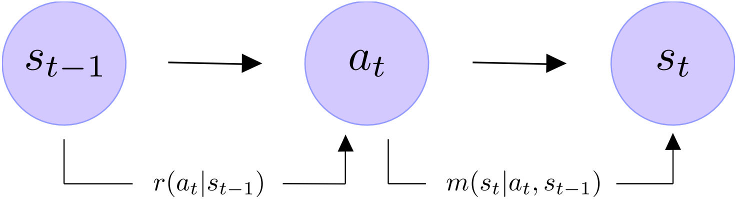


Figure: System evolution in time.

- time step $t \in \mathcal{S}_t$, where $|\mathcal{S}_t| = N$
- states $s_t, s_{t-1} \in \mathcal{S}_s$
- action $a_t \in \mathcal{S}_a$
- decision strategy $r(a_t|s_{t-1})$, $t \in \mathcal{S}_t$
- system model $m(s_t|a_t, s_{t-1})$

State Transitions

- unknown state transitions

$$m(s_t|a_t, s_{t-1}) \quad \rightarrow \quad m(s_t|a_t, s_{t-1}, \theta)$$

- system evolution history

$$H_t = (s_t, a_t, s_{t-1}, a_{t-1}, ..., s_1, a_1, s_0)$$

- parameter density $p(\theta|H_t) \quad \rightarrow \quad$ adaptivity

$$p(\theta|H_t) = \frac{m(s_t|a_t, s_{t-1}, \theta)p(\theta|H_{t-1})}{\int_{\mathcal{S}_\theta} m(s_t|a_t, s_{t-1}, \theta)p(\theta|H_{t-1})d\theta}$$

Forgetting

- parameters varying in time

$$p(\theta|H_t) \quad \rightarrow \quad p(\theta_t|H_t)$$

- forgetting factor λ_t
- set of hypotheses:

$$p(\theta_{t+1}|H_t) = \begin{cases} p(\theta_t|H_t) & \text{with probability } \lambda_t \\ p(\theta_t|H_0) & \text{with probability } 1-\lambda_t \end{cases}$$

- combination for $\lambda_t \in [0,1]$, see [3]

$$p(\theta_{t+1}|H_t) \propto [p(\theta|H_t)]^{\lambda_t} [p(\theta|H_0)]^{1-\lambda_t}$$

Fully Probabilistic Design

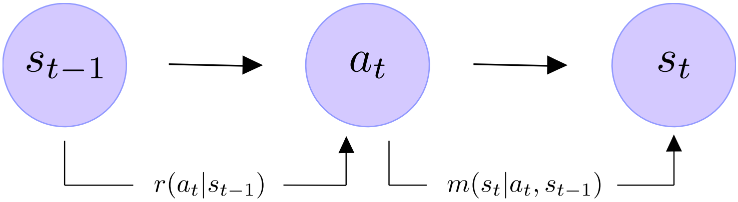


Figure: System evolution in time.

- process optimization
- decision strategy choice [1]
- system exploration

Optimal Decision Strategy

- optimized decision rules

$$r^o(a_t|s_{t-1}) = \frac{r^i(a_t|s_{t-1})\exp(-d(a_t, s_{t-1}))}{h(s_{t-1})}$$

- auxiliary functions

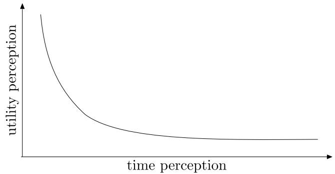
$$h(s_{t-1}) \equiv \int_{\mathcal{S}_a} r^i(a_t|s_{t-1})\exp\left(-d(a_t, s_{t-1})\right) da_t,$$

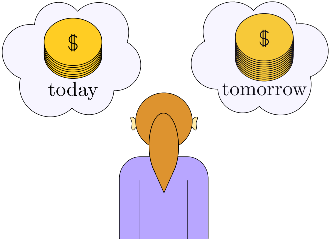
$$d(a_t, s_{t-1}) \equiv \int_{\mathcal{S}_s} m(s_t|a_t, s_{t-1}) \ln\left(\frac{m(s_t|a_t, s_{t-1})}{h(s_t)m^i(s_t|a_t, s_{t-1})}\right) ds_t,$$

$$h(s_N) \equiv 1$$

Discounting

- discounting = current utility of future gains (losses)
- commonly utilized in economy fields





Optimal Decision Strategy in Discounted Case

- optimization $\iff$ standard fully probabilistic design (no discounting)

$$r^o(a_t|s_{t-1}) = \frac{r^i(a_t|s_{t-1})\exp(-d(a_t, s_{t-1}))}{h^i(s_{t-1})}$$

- t-th step omitted from the optimization (discounting)

$$a^o(s_{t-1}) \in \operatorname*{argmin}_{a_t \in \mathcal{S}_a} \int_{s_t \in \mathcal{S}_s} m(s_t|a_t, s_{t-1}) \ln\left(\frac{1}{h(s_t)}\right) ds_t$$

Certainty-Equivalence Strategy

- growing dimensionality of the problem
- formation of hyperstates
- makes use of parameter point estimates [2]

$$m(s_t|a_t, s_{t-1}, \hat{\theta})$$

Simulation Setup

- $\mathcal{S}_s = \{1, 2, 3\}$, $\mathcal{S}_a = \{1, 2\}$
- preferences: state 3, action 2
- 200 time steps
- averaging over 50 seed values
- modeling error

Results

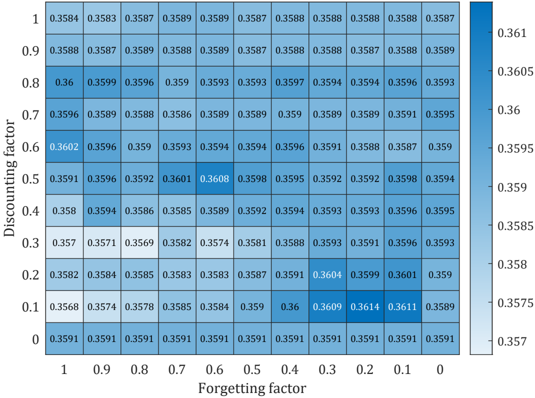


Figure: Occurrence of the preferred state 3 in percents.

- optimal factor combination found near the diagonal
- improvement is not significant

Summary

- Main contribution of the work:
  - combining discounting and forgetting into one task
- Possible focus of the following research:
  - discounting factor choice
  - continuous case → regression

Bibliography

- M. Kárný.
  Towards fully probabilistic control design.
  Automatica, 32:1719–1722, 1996.

- M. Kárný, A. Halousková, J. Böhm, R. Kulhavý, and P. Nedoma.
  Design of linear quadratic adaptive control: theory and algorithms for practice.
  Kybernetika, 21 (1a):3–96, 1995.

- R. Kulhavý and M. B. Zarrop.
  On a general concept of forgetting.
  International Journal of Control, 58(4):905–924, 1993.