Magic number five: The breadth—depth dilemma in accumulator and tree-like models of decision making

Jorge Ramírez-Ruiz*, Chiara Mastrogiuseppe* and Rubén Moreno-Bote

Universitat Pompeu Fabra Barcelona

CBC Center for Brain and Cognition

GOBIERNO DE ESPAÑA · MINISTERIO DE CIENCIA E INNOVACIÓN · AGENCIA ESTATAL DE INVESTIGACIÓN

Generalitat de Catalunya

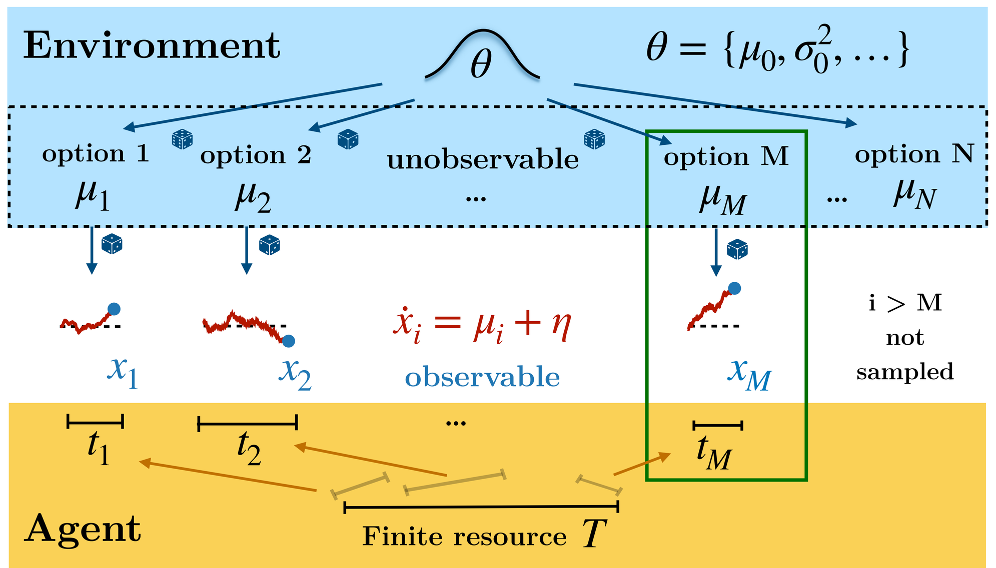
Unió Europea
Fons Social Europeu
L'FSE inverteix en el teu futur

Motivation

- We study the problem of **allocating finite** sampling **resources** to determine the best of several options.
- What does it mean to decide **optimally** under resource constraints? How does the environment contribute to the optimality of the solutions?
- Why is it good to ignore many options in many cases?
- Thus far, research has considered low numbers of available options and is not explicit about limitation of resources.
- We study the **optimal allocation policy** in two different models: an **accumulator** and a **tree-like** model.

Accumulator Model

- Environment produces many options, agent is familiar with environment, but ignores the true value of each option.
- After allocation, agent obtains final evidence and chooses the option with highest inferred drift (green box).
- Expected utility for allocation is over possible evidences:

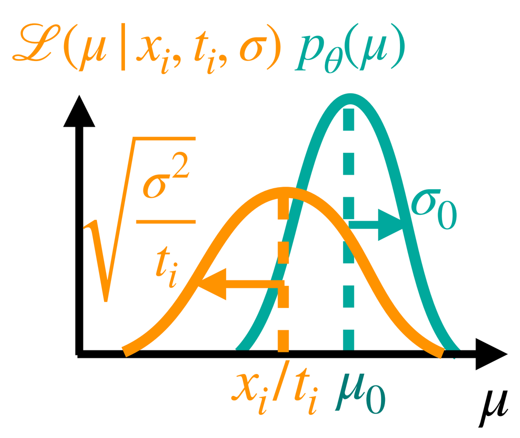


$$\hat{U}(M, \mathbf{t}) \equiv \mathbb{E}\left[\max_{i \le M} \hat{\mu}_i \,|\, \mathbf{t}\right] = \int dx_1 \dots dx_M p(x_1, \dots, x_M \,|\, \mathbf{t}) \max_i \hat{\mu}_i(x_i, t_i)$$

Capacity and optimal policies

Sampling capacity of the agent is the ratio between precision of the observations and precision of the prior:

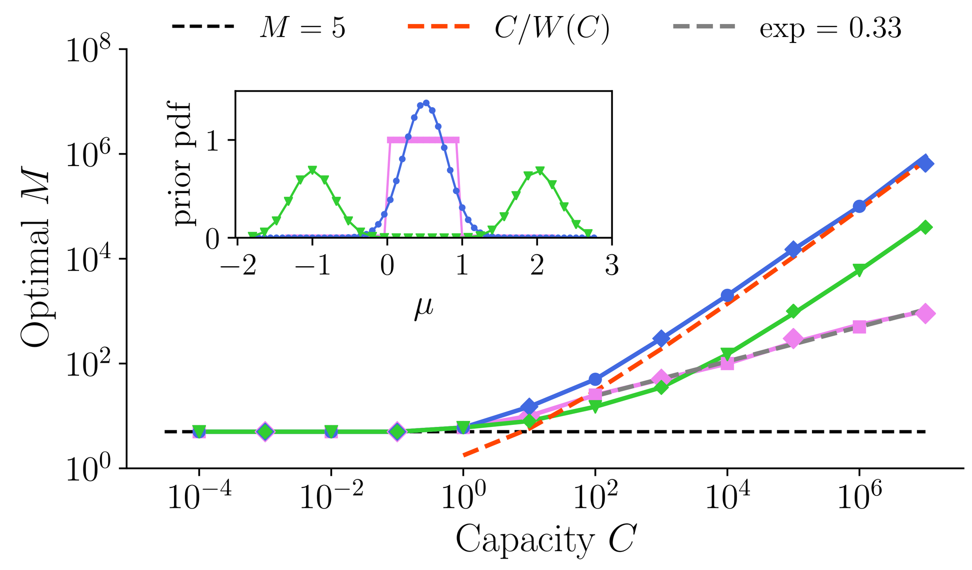
$$C = \sum_{i=1}^{M} \frac{\sigma_0^2}{\sigma^2/t_i} = \frac{\sigma_0^2}{\sigma^2} T$$



For even allocations and small capacity,

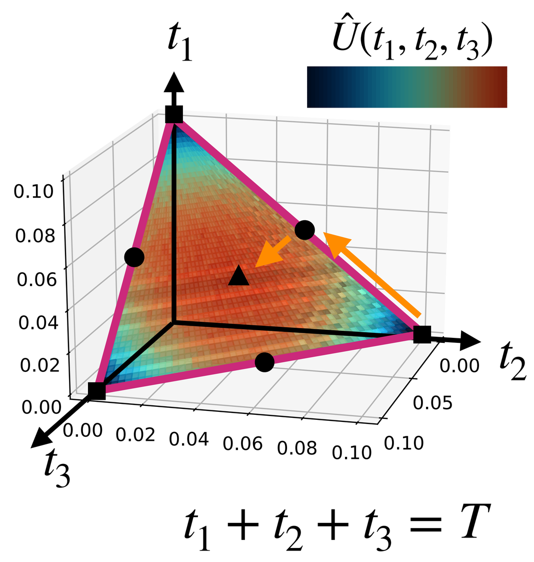
$$\hat{U}(M, C) \approx \mu_0 + \sigma_0 \sqrt{\frac{C}{2\pi}} \left[\sqrt{M} \int_{-\infty}^{\infty} dz\, z \exp(-\frac{z^2}{2}) \left(\frac{1}{2} + \frac{1}{2}\mathrm{erf}\left(\frac{z}{\sqrt{2}}\right)\right)^{M-1}\right]$$

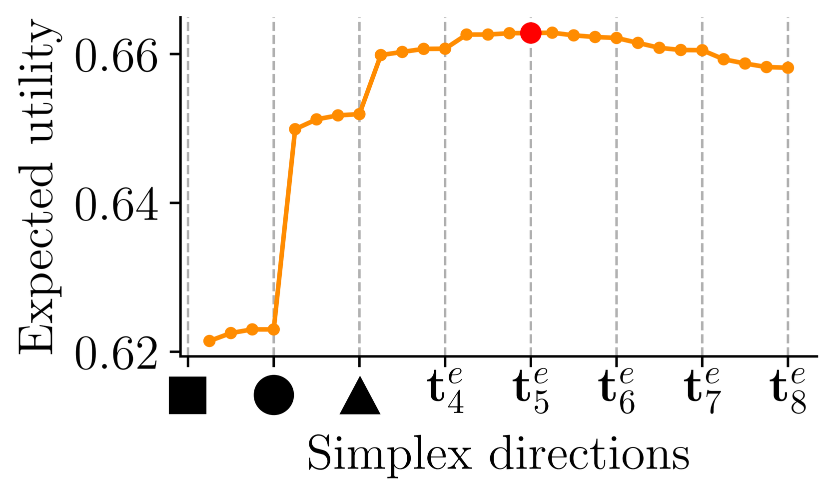
has a maximum at $M^* = 5$.



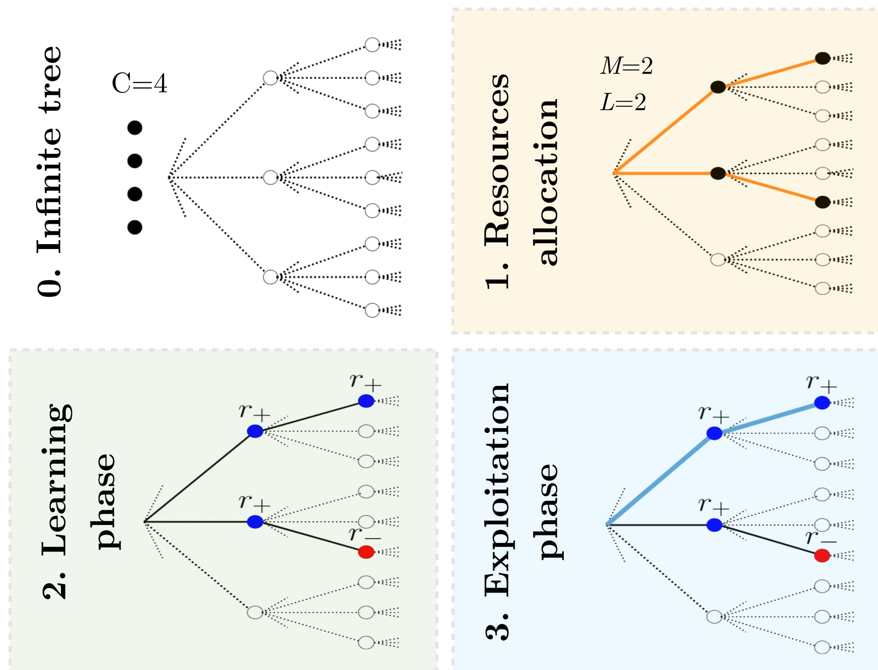
Even allocation is optimal

Since options are a priori indistinguishable, allocating capacity evenly between M^* options ($t^e_{M^*,i} = T/M^*$ for $i = 1, \dots, M^*$) is optimal, which we verify with numerical optimization.



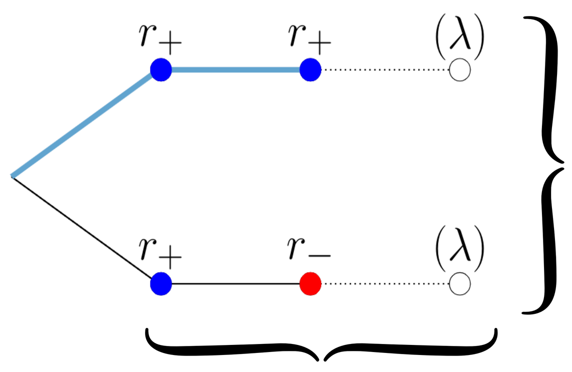


Tree Model



- The agent allocates finite given samples C over an infinite decision tree by choosing the number of M non-branching paths of equal length L, as deep as the resources allow (yellow box).
- Remaining samples are equally distributed into another level $L+1$ with probability λ.
- Agent discovers the rewards in each node (green box) and chooses the path with highest accumulated reward (blue box).
- r_+ is collected with probability p with value s.t. $\mathbb{E}[r] = 0$

Diffusion—Maximization Algorithm



2. Maximization step: Probability of the cumulative reward over the optimal path J

$$P(J_{L+1} = j) = \left(P(R_{L+1} \le j)\right)^M - \left(P(R_{L+1} \le j - 1)\right)^M$$

1. Diffusion step: Probability of the cumulative reward in a single path R

$$P(R_L) \longrightarrow P(R_{L+1})$$

$$R_L \sim Bin(i; L, p)$$

$$R_L = \sum_{j=1}^{L} r_j = i \cdot r_+ + (L - i) \cdot r_- = k(i)$$

$$P(R_{L+1} = k(i) + r_+) = \lambda p P(R_L = k(i))$$

$$P(R_{L+1} = k(i)) = (1 - \lambda) P(R_L = k(i))$$

$$P(R_{L+1} = k(i) + r_-) = \lambda (1 - p) P(R_L = k(i))$$

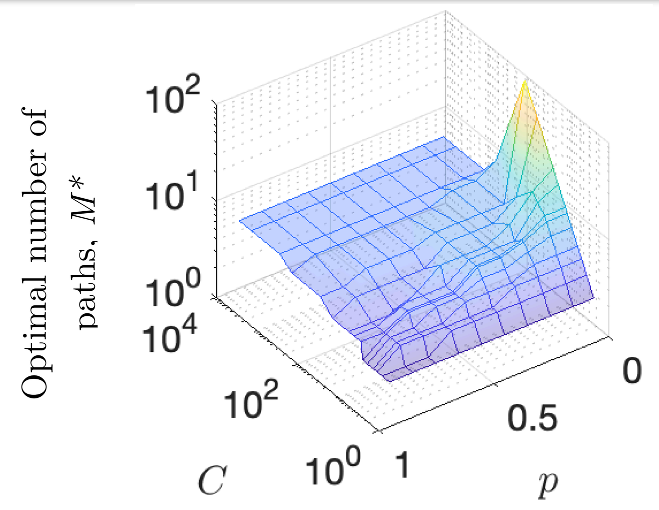
Exact algorithm: $R_L \sim Bin(i; L, p)$ $\xrightarrow{L \gg 1}$ **Gaussian approximation**: $R_{L+1} \sim \mathcal{N}(x; \mu, \sigma^2)$

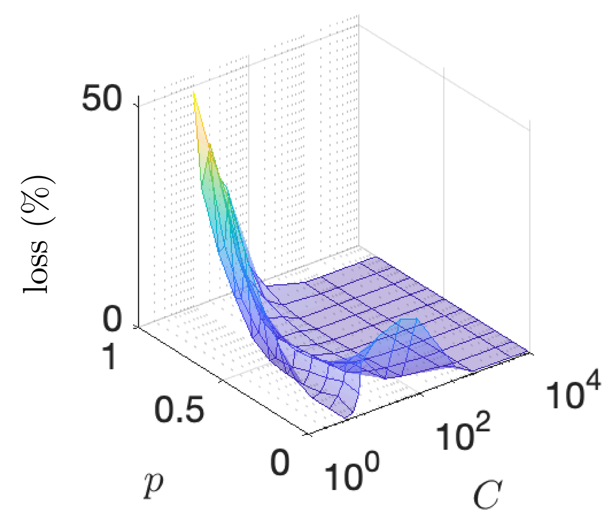
- ✓ For L large enough
- ✓ Continuous variables
- ✓ Continuous maximization step

Deep allocations are close to optimal

Value as the expected cumulative reward over the optimal path, $V_{C,M} = \mathbb{E}[J_{L+1}]$

The optimal policy is the number of paths that maximises the value, $M^* = \arg\max V_{C,M}$





- Large regions of the plane are dominated by deep allocations $M^* \sim 5$
- Breadth dominates for small values of p in a small range of capacity C: $M^* \sim 5$

Little loss occurs by using always $M^* \sim 5$

Take-away messages

- Considering **five options** is optimal for a wide range of capacities and environments, in both models.
- When a **large** number of resources are available:
  - -in the accumulator model the optimal number of sampled options grows sub-linearly with capacity, i.e. emphasis of **depth over breadth;**
  - -in the tree model it is optimal to consider **five** options regardless of the richness of the environment p.
- When dealing with a **small** amount of resources:
  - -in the accumulator model it is optimal to consider **five** options regardless of the prior;
  - -in the tree model there are regions dominated by **breadth,** although in those cases **little loss** would occur by choosing a different policy.