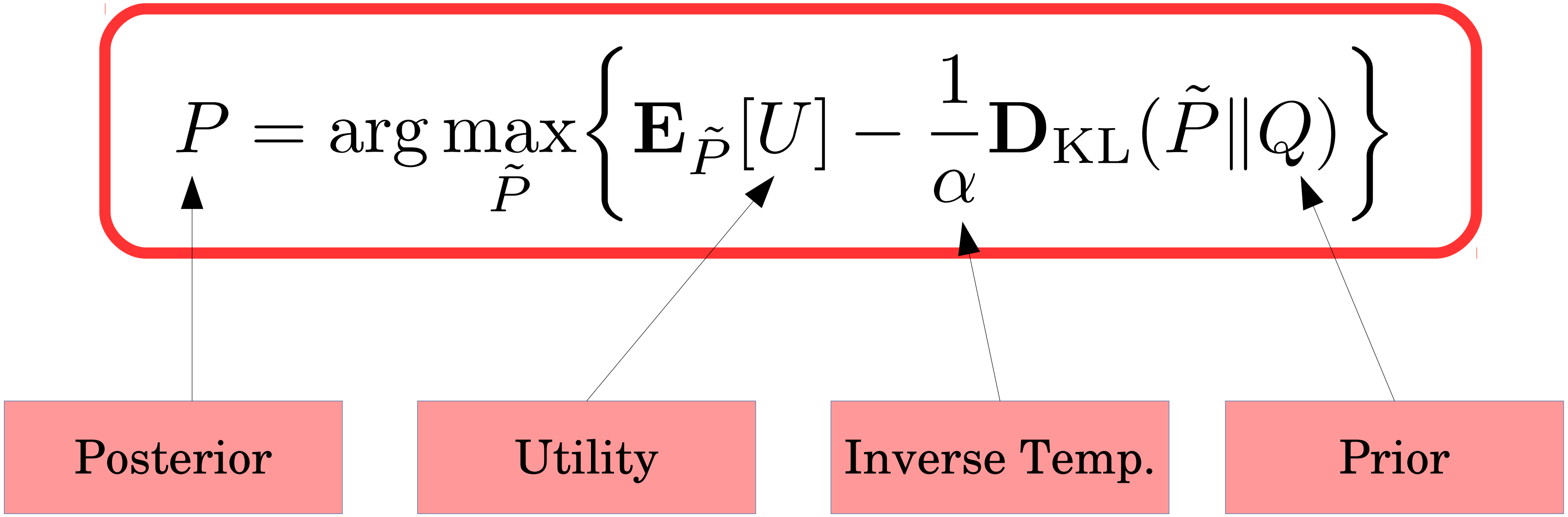
Information-Theoretic Bounded Rationality

Pedro A. Ortega
University of Pennsylvania

Goal of this Talk

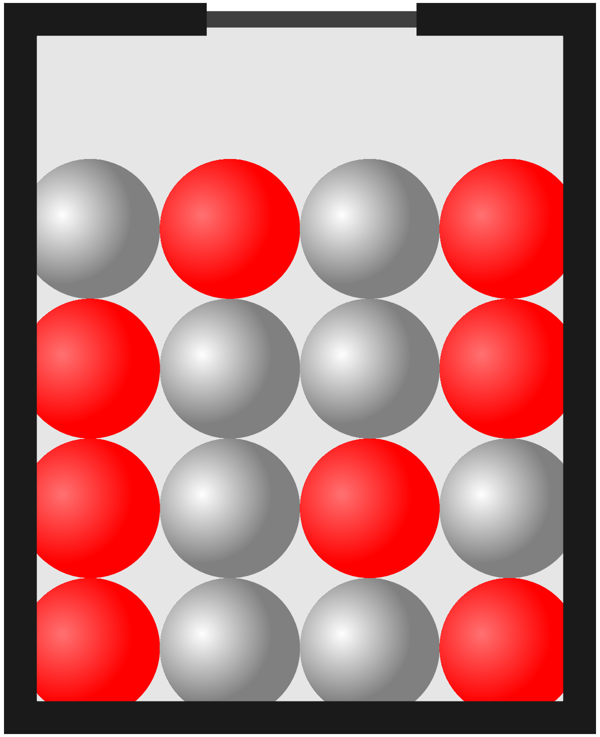
In large-scale decision problems, agents choose as



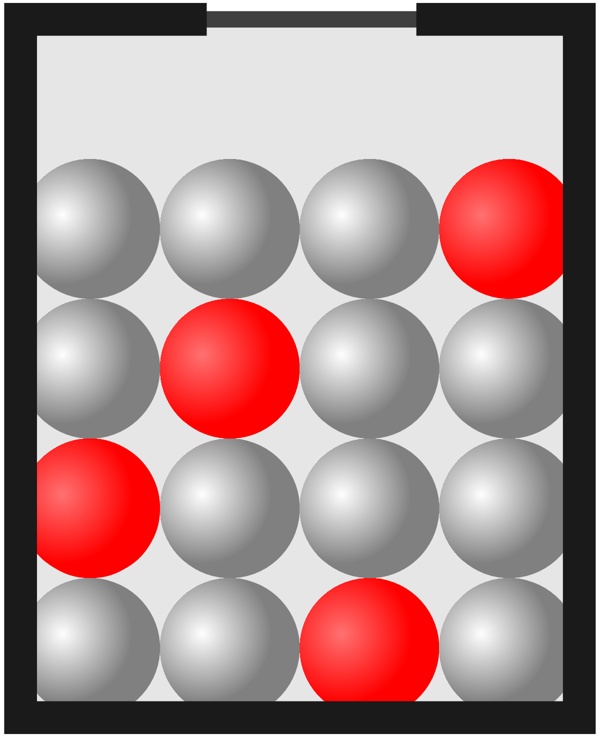
$$P = \arg\max_{\tilde{P}} \left\{ \mathbf{E}_{\tilde{P}}[U] - \frac{1}{\alpha} \mathbf{D}_{\mathrm{KL}}(\tilde{P} \| Q) \right\}$$

Why is this so? And how is this used?

Pick a box.

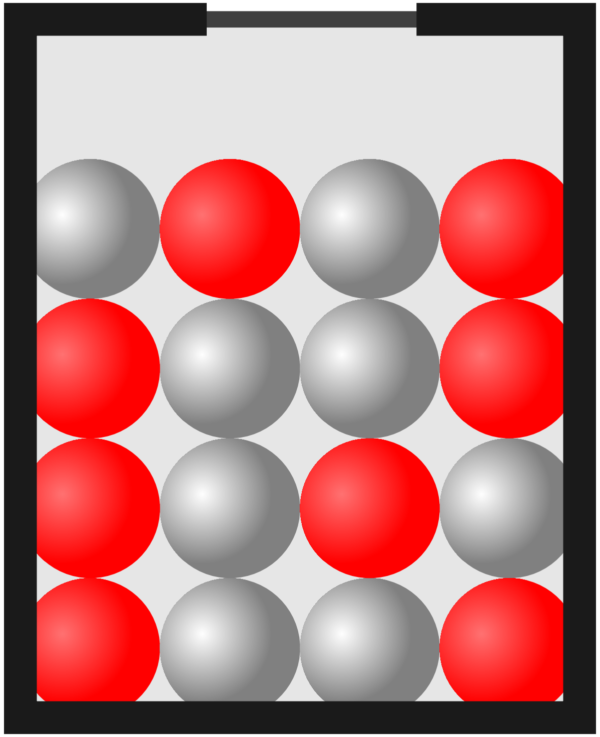


50/50



25/75

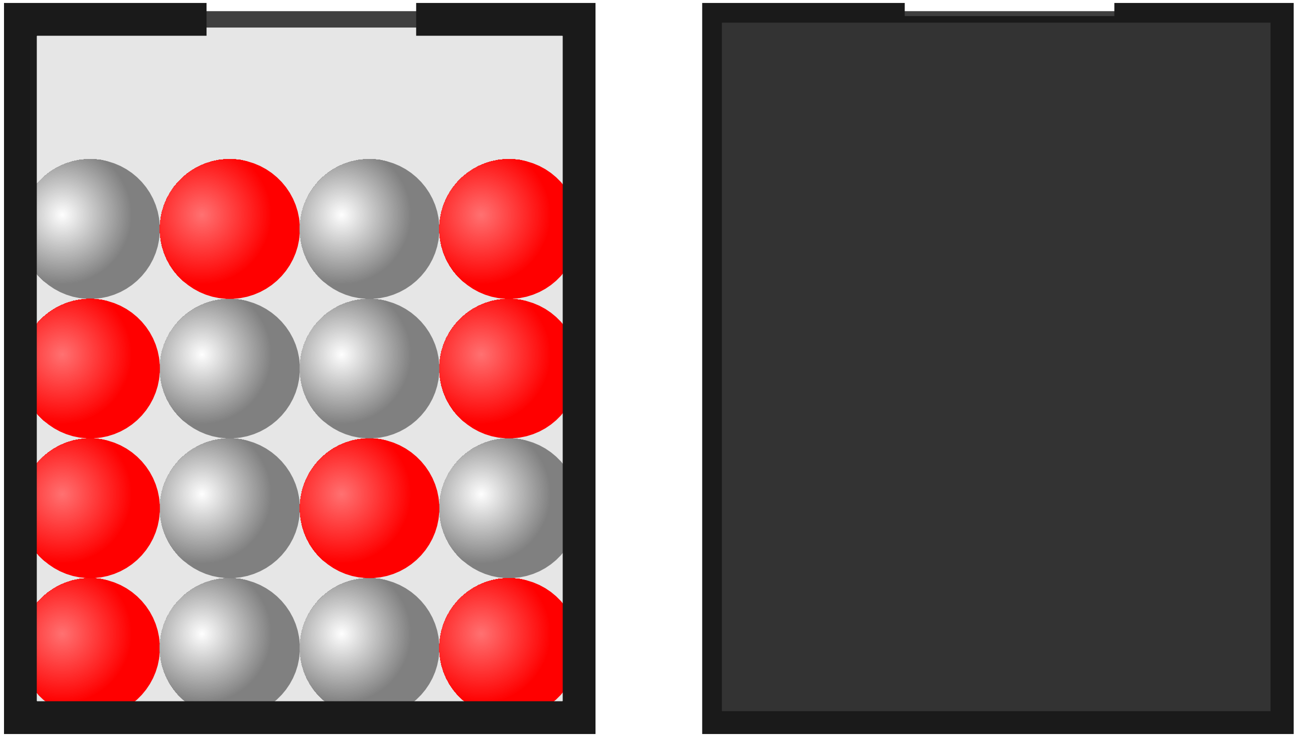
Pick a box.



50/50

?

Pick a box.

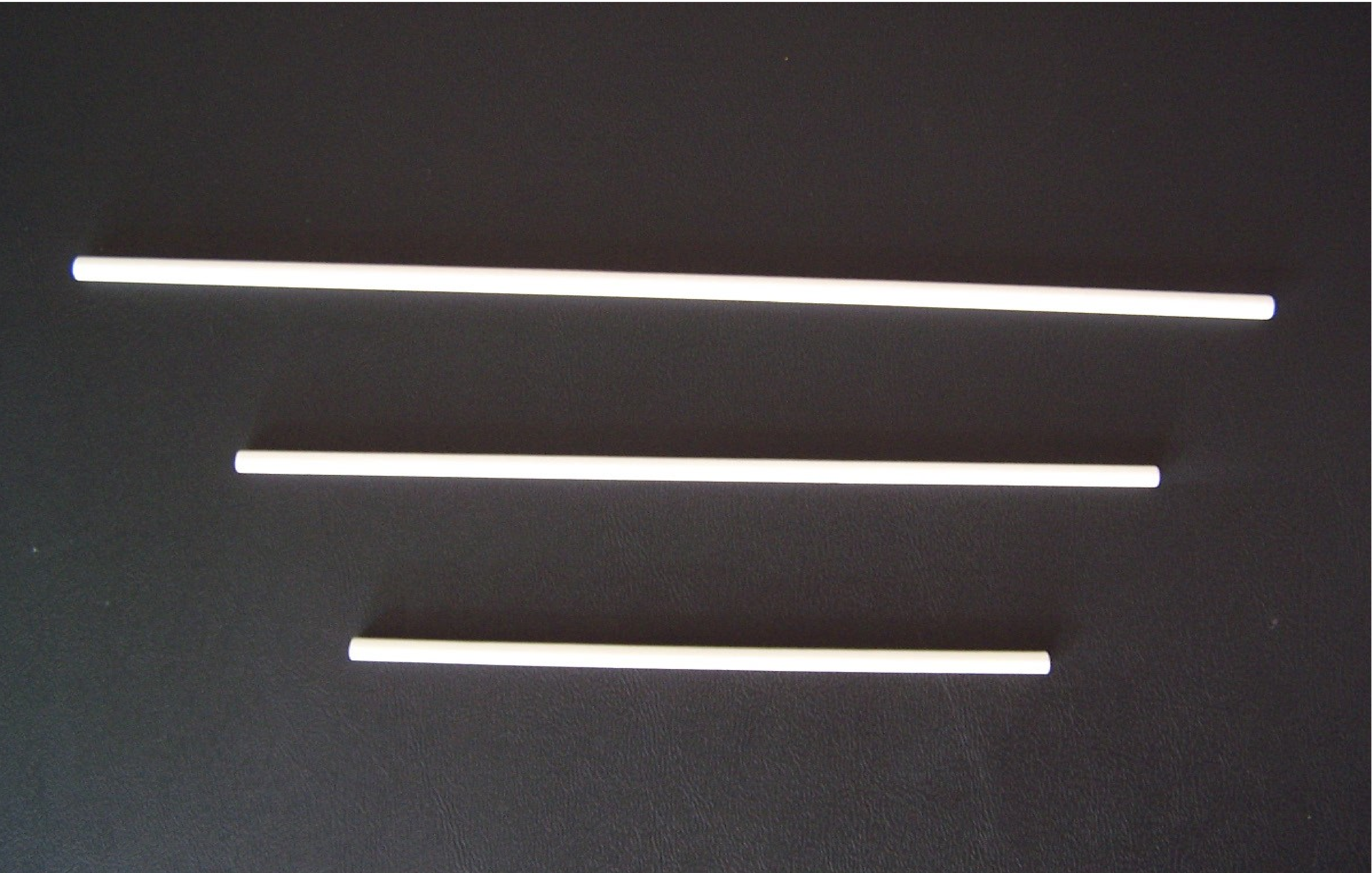


50/50

?

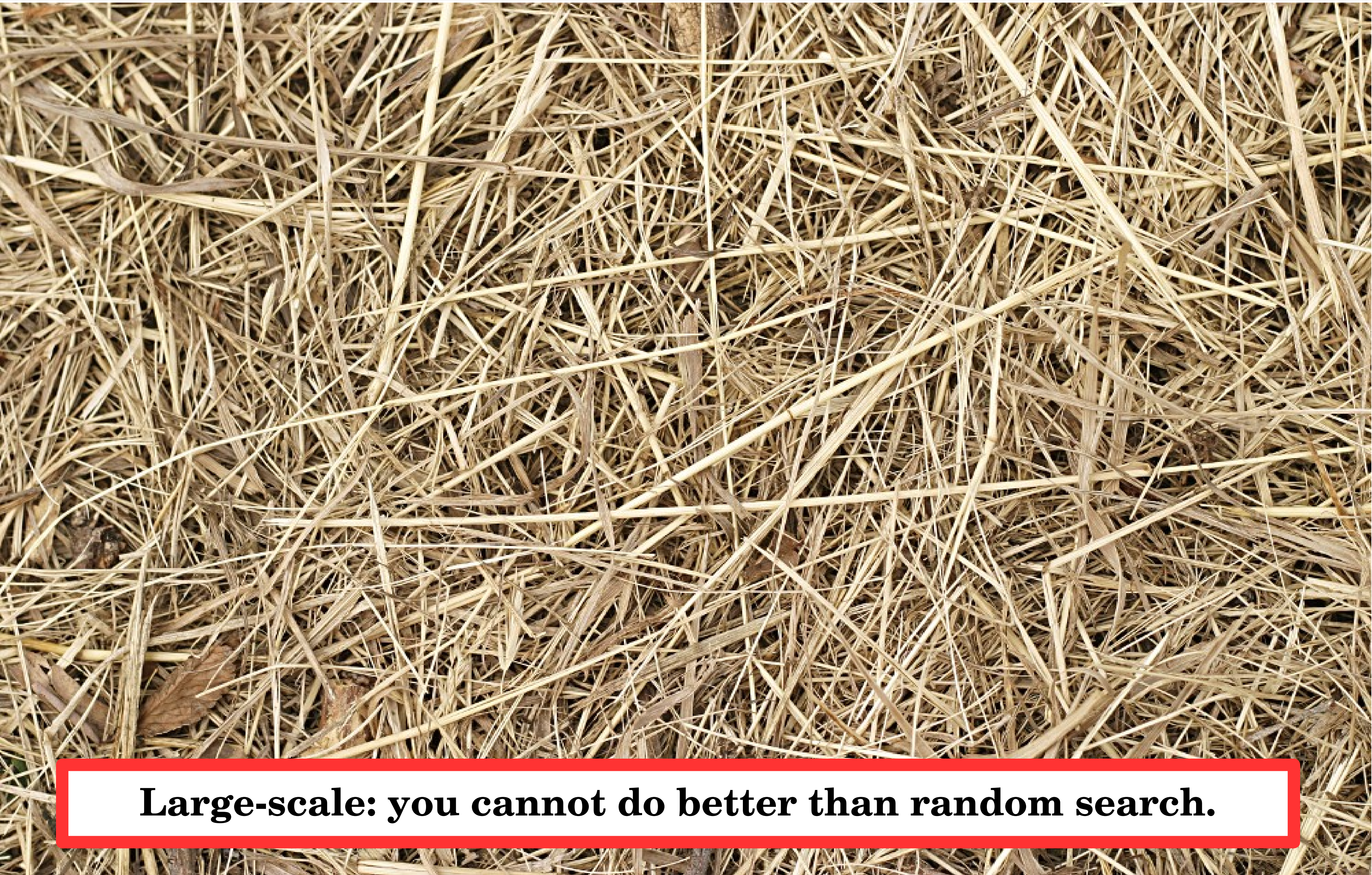
Unknown probabilities affect net value.

Pick the largest.



Pick the largest.

Pick the largest.



Large-scale: you cannot do better than random search.

Preliminaries

History, variational principles and more

An active area of research...

Polani

ermodynamics of Prediction

Still

Duality Control-Estimation

Mitter

Gomez

Marinacci

Regularisation in Control

Rational Inattention

Linearly-Solvable Control

Peters

Variational Preferences

Maccheroni

Todorov

Wiegerink

KL-Control Theory

Neumann

Theodorou

Friston

Tishby

Rustichini

Active Inference

Palfrey

Bierkens

Kappen

Information Flow in Perception-Ac

Robust Decision Making

Fleming

Guerra

McKelvey

Sims

Braun

Quantum Mechanics for Control

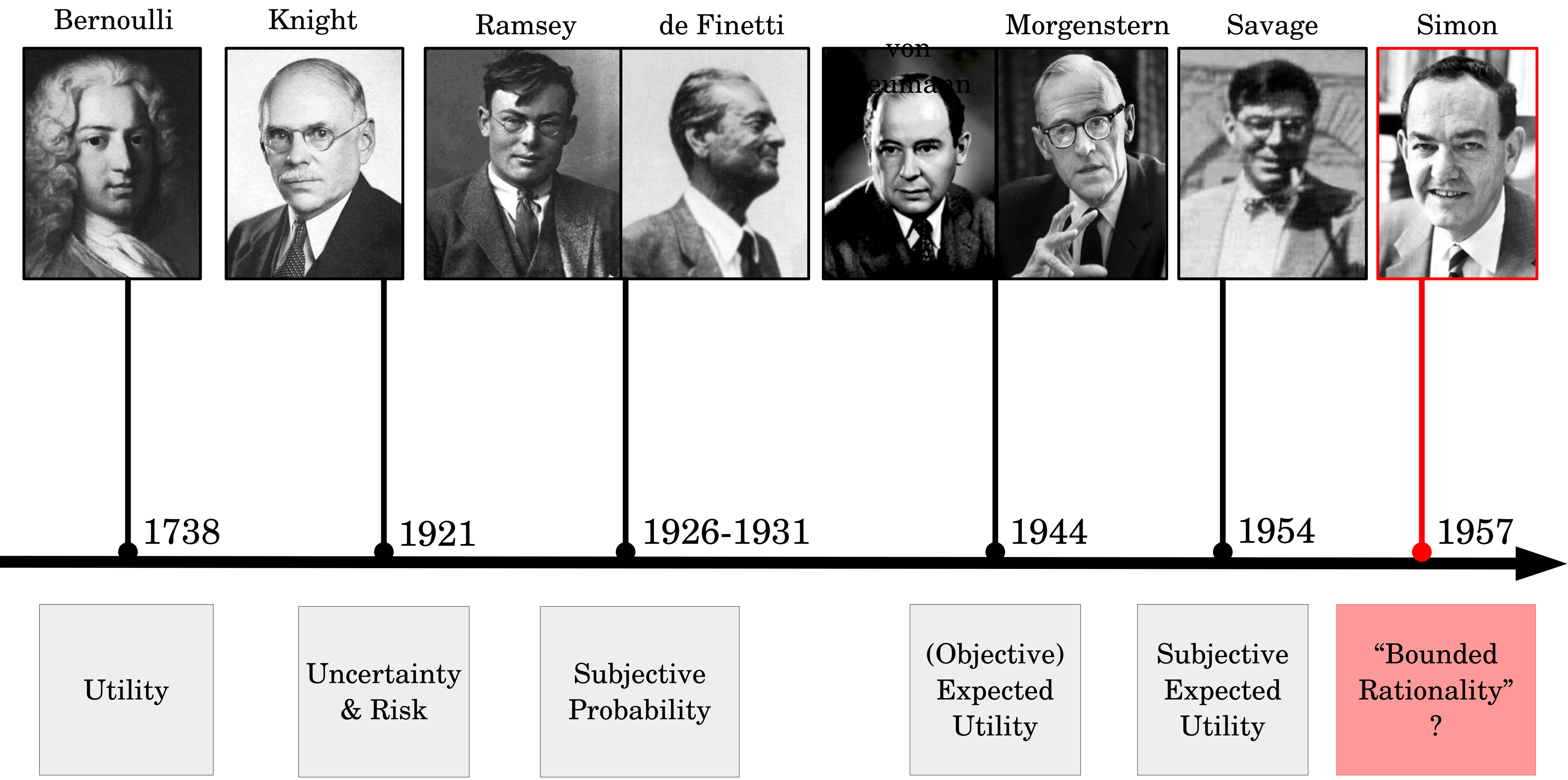
Orteg

Wolpert

Quantal Response Equilibria

Bounded Rationality

History of Decision Theory



Bernoulli

Knight

Ramsey

de Finetti

von
Neumann

Morgenstern

Savage

Simon

1738

1921

1926-1931

1944

1954

1957

Utility

Uncertainty
& Risk

Subjective
Probability

(Objective)
Expected
Utility

Subjective
Expected
Utility

“Bounded
Rationality”
?

New theory of bounded-rational decisions.

What is boundedness?

NO

- Suboptimal
- Heuristics
- Approximation
- Meta-reasoning

YES

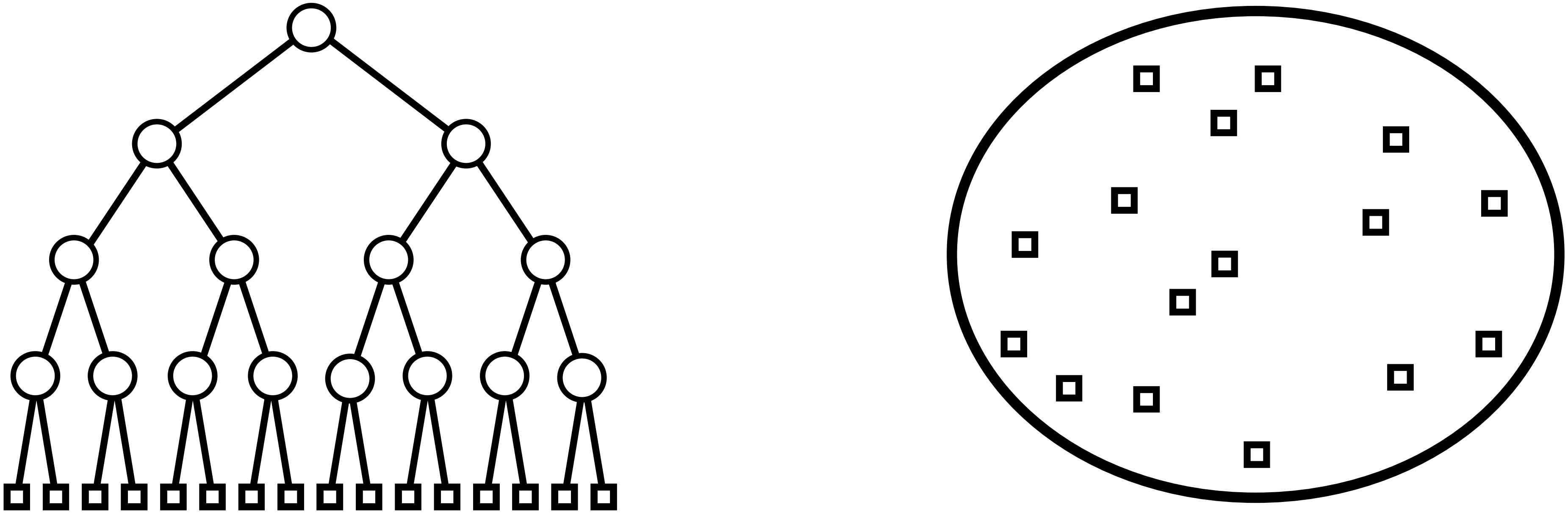
- Intractability
- Model Uncertainty
- Risk Sensitivity
- ToM

Bounded = Subject to information constraints.

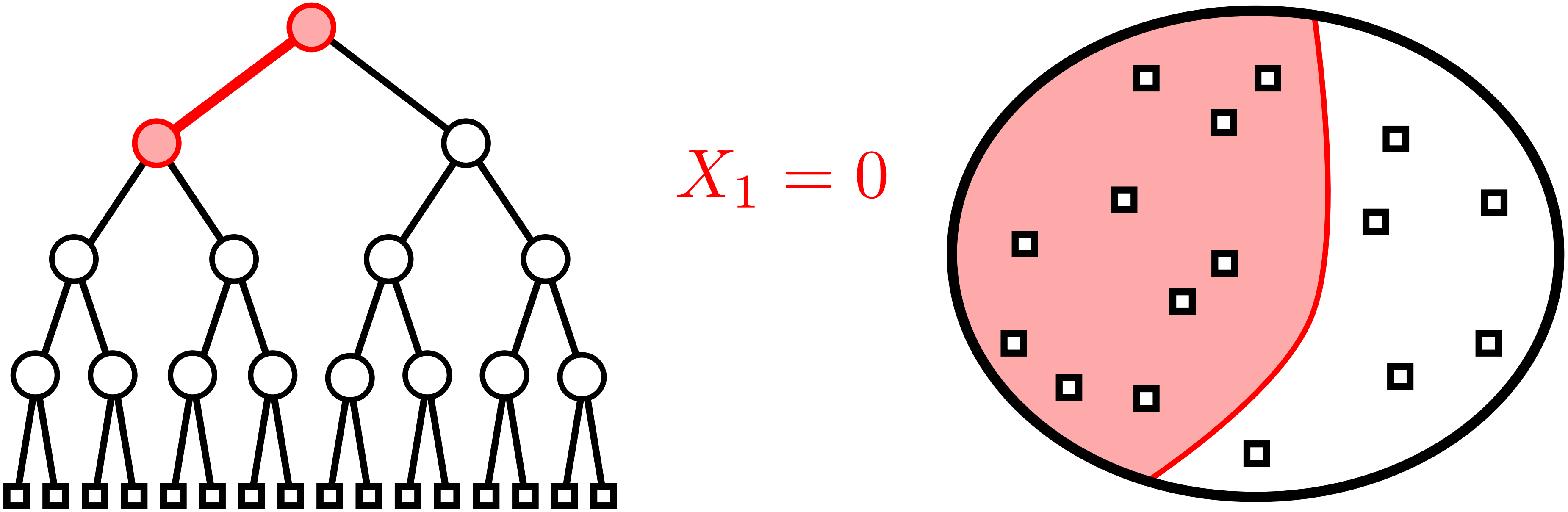
Bounded-Rational Decisions

Assumptions, derivation and main properties

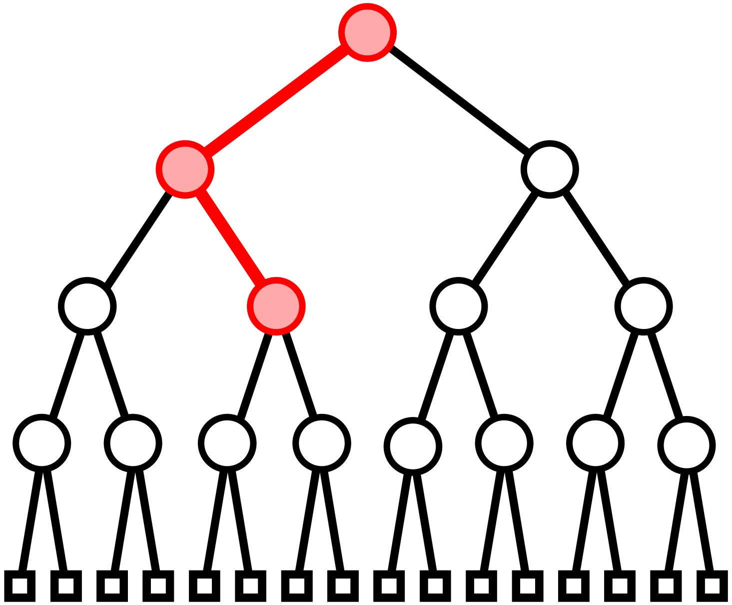
Planning and Searching



Planning and Searching

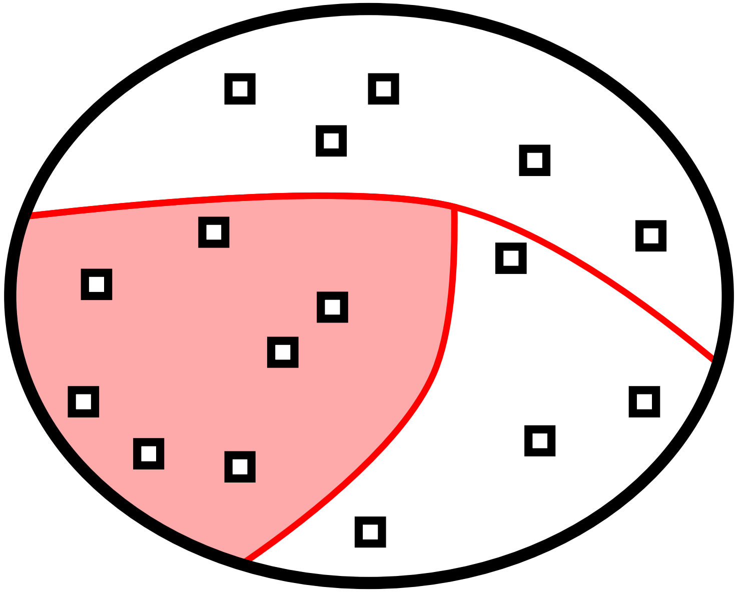


Planning and Searching

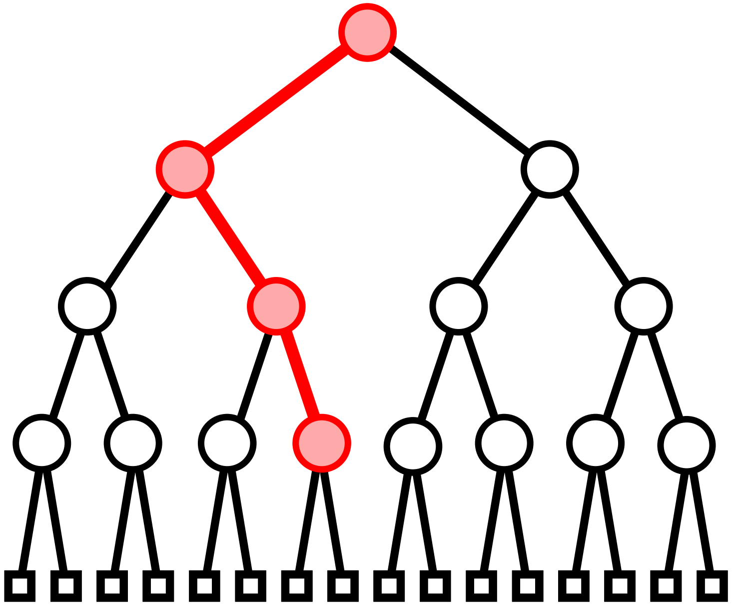


$X_1 = 0$

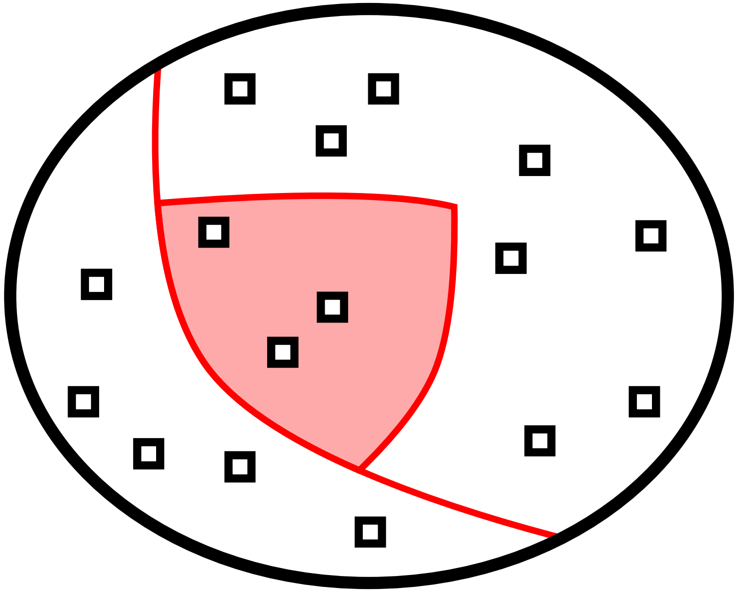
$X_2 = 1$



Planning and Searching



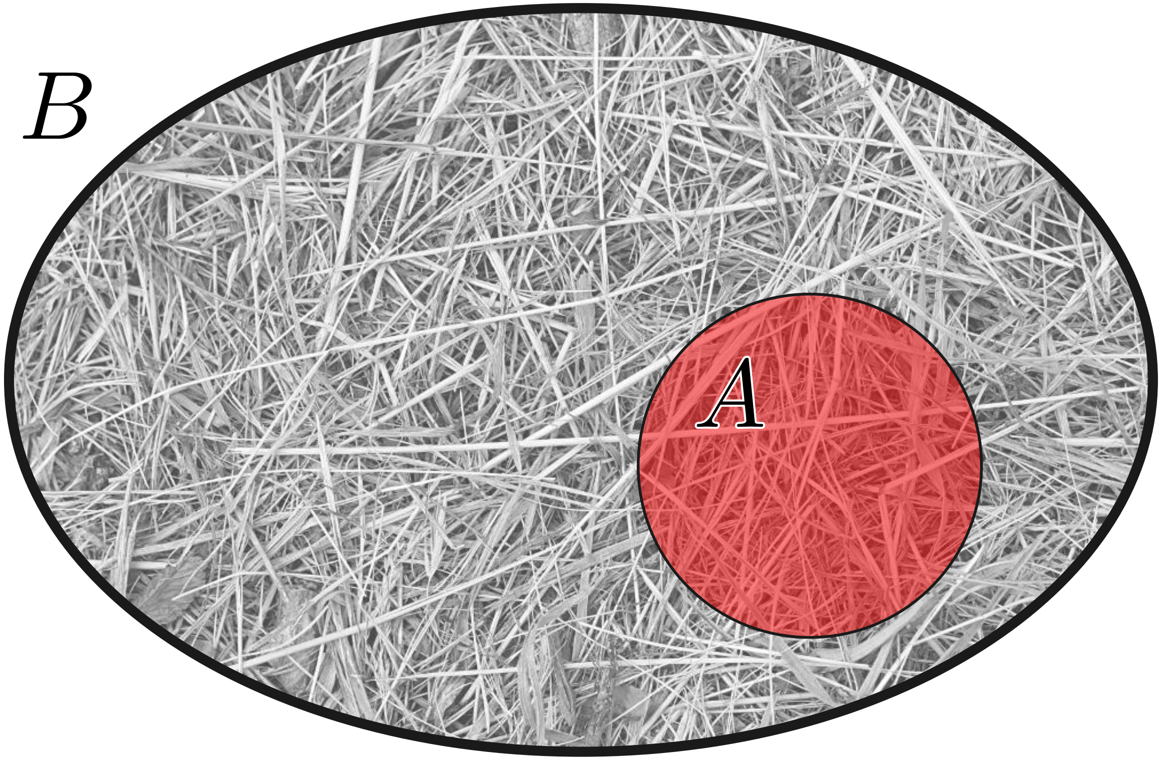
$$X_1 = 0$$
$$X_2 = 1$$
$$X_3 = 1$$



Planning is searching for a “good” subset.

“Straw Pile” Assumptions

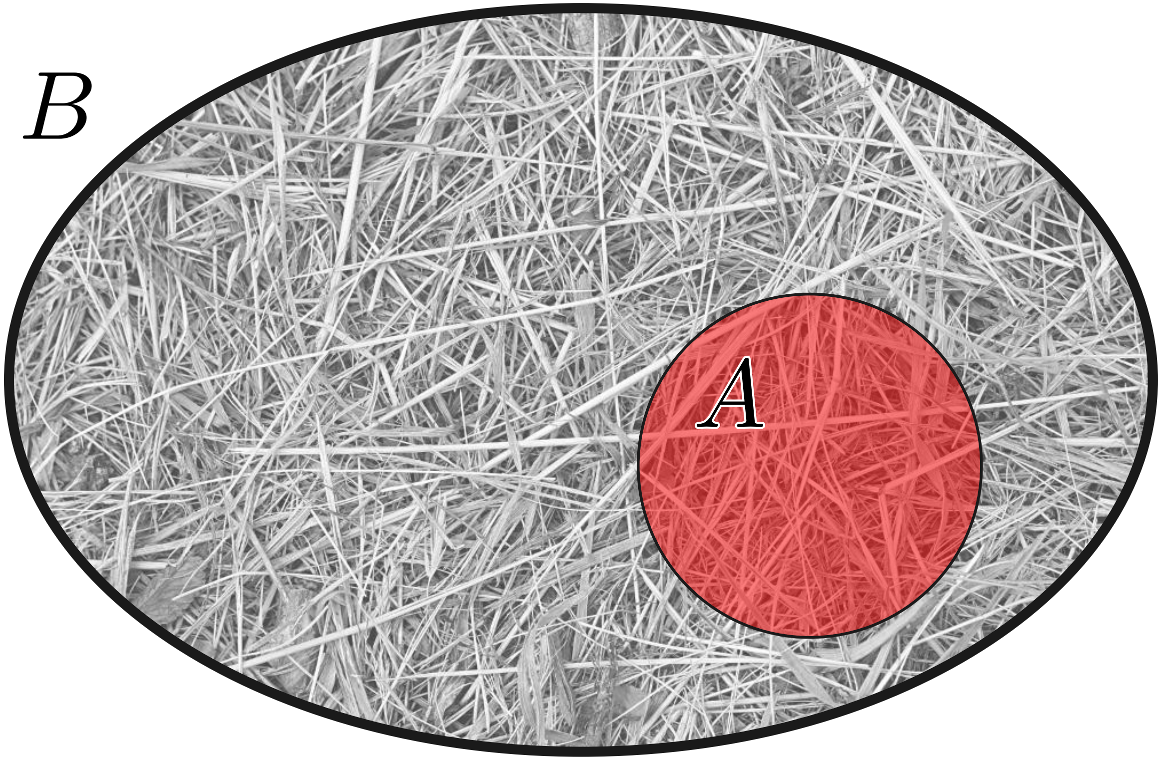
Primitive operation: search A given B.



- Very large
- Negligible probability
- Irreducible
- Parallelisable

“Straw Pile” Assumptions

Primitive operation: search A given B.



- Very large
- Negligible probability
- Irreducible
- Parallelisable

Golden Rule:
No Integration!

Large-scale planning is searching a set inside a pile of straw.

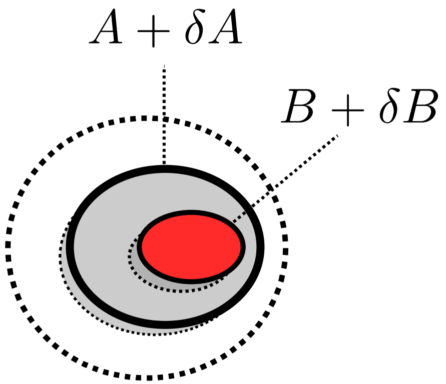
Search Costs

$$\boxed{\mathbf{C}(A|B) = -\frac{1}{\alpha}\log \mathbf{Pr}(A|B) = \frac{1}{\alpha}\log \mathbf{E}[N]}$$

Search Costs

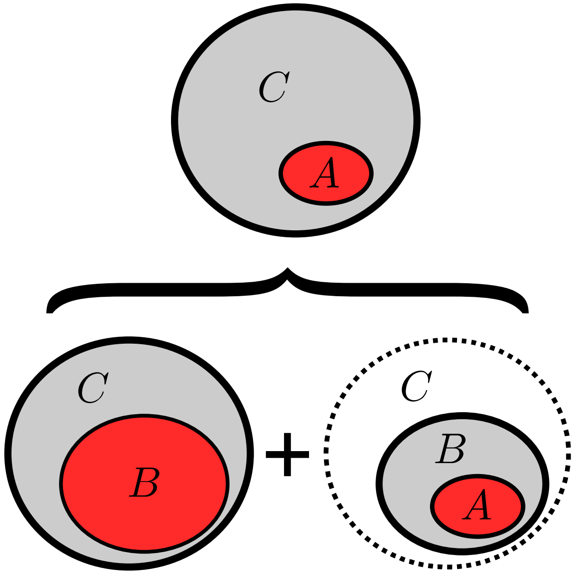
$$\mathbf{C}(A|B) = -\frac{1}{\alpha}\log \mathbf{Pr}(A|B) = \frac{1}{\alpha}\log \mathbf{E}[N]$$

Continuity



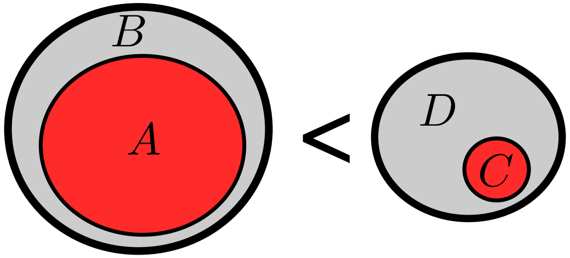
$$\mathbf{C}(A|B) = f \circ \mathbf{Pr}(A|B)$$

Additivity



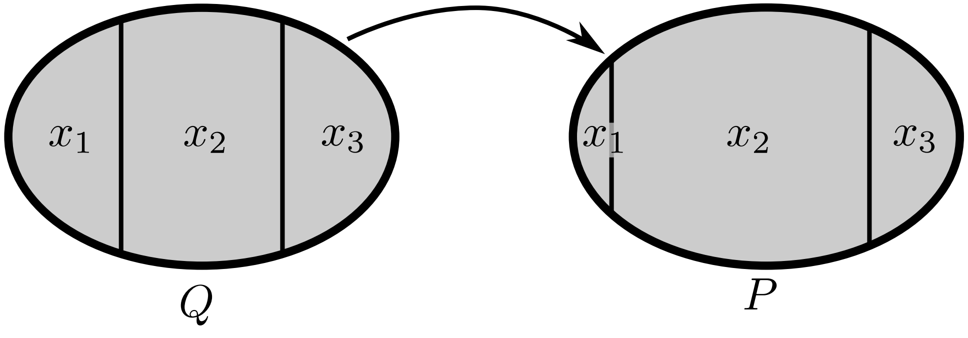
$$\mathbf{C}(A \cap B|C) = \\ \mathbf{C}(B|C) + \mathbf{C}(A|B \cap C)$$

Monotonicity

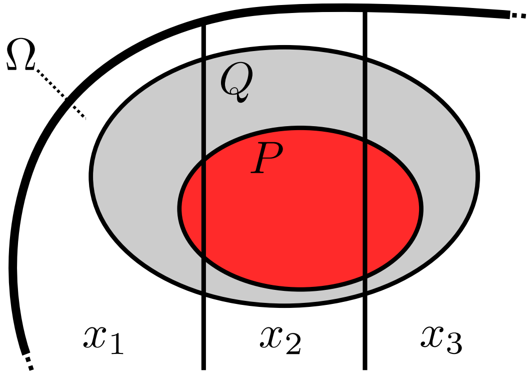


$$\mathbf{Pr}(A|B) > \mathbf{Pr}(C|D) \\ \Leftrightarrow \mathbf{C}(A|B) < \mathbf{C}(C|D)$$

Cost of Probabilistic Choice



a)



b)

$$\mathbf{C}(P|Q) = \sum_x P(x)\mathbf{C}(x \cap P | x \cap Q) + \frac{1}{\alpha} \sum_x P(x) \log \frac{P(x)}{Q(x)}$$

$Q(\cdot) = \mathbf{Pr}(\cdot|Q) \quad P(\cdot) = \mathbf{Pr}(\cdot|P)$

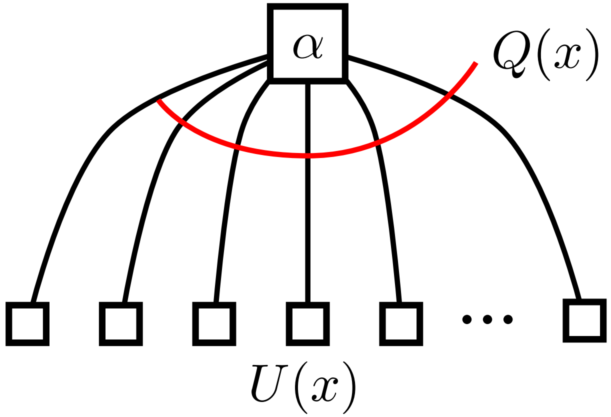
Total Cost = Expected Cost + Waste

Bounded-Rational Decision Making

Deliberation transforms prior Q into posterior P.

Agent extremises the **free energy functional**:

$$\boxed{F[\tilde{P}] = \mathbf{E}_{\tilde{P}}[U] - \frac{1}{\alpha}\mathbf{D}_{\mathrm{KL}}(\tilde{P}\|Q)}$$



Trades off utility & information costs.

Models:

- large-scale choice spaces,
- trust & risk sensitivity,
- and other **information constraints**.

Optimal Choice & Certainty-Equivalent

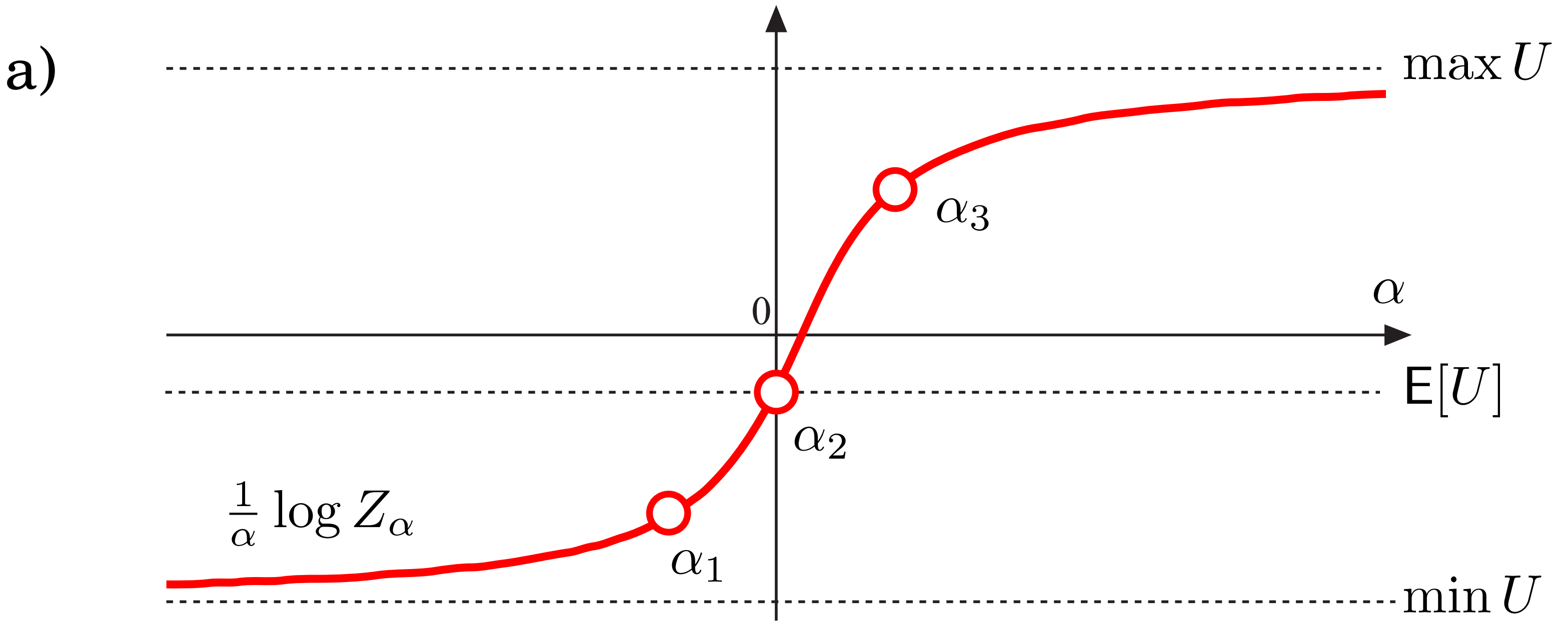
Posterior (optimal choice distribution):

$$P(x) = \frac{1}{Z} \cdot Q(x) \cdot e^{\alpha U(x)}$$

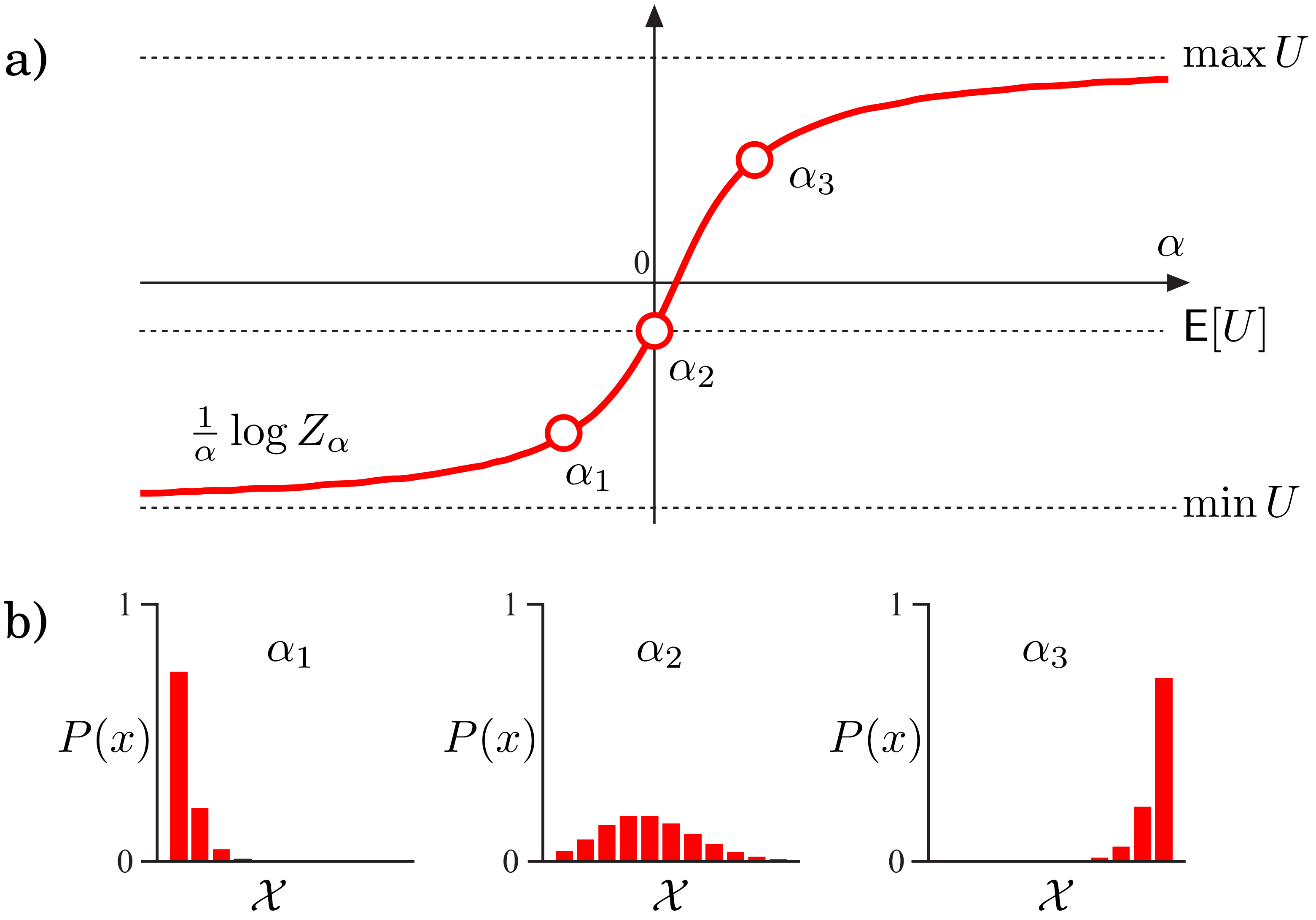
Certainty-Equivalent (value):

$$F[P] = \frac{1}{\alpha} \log Z_\alpha$$

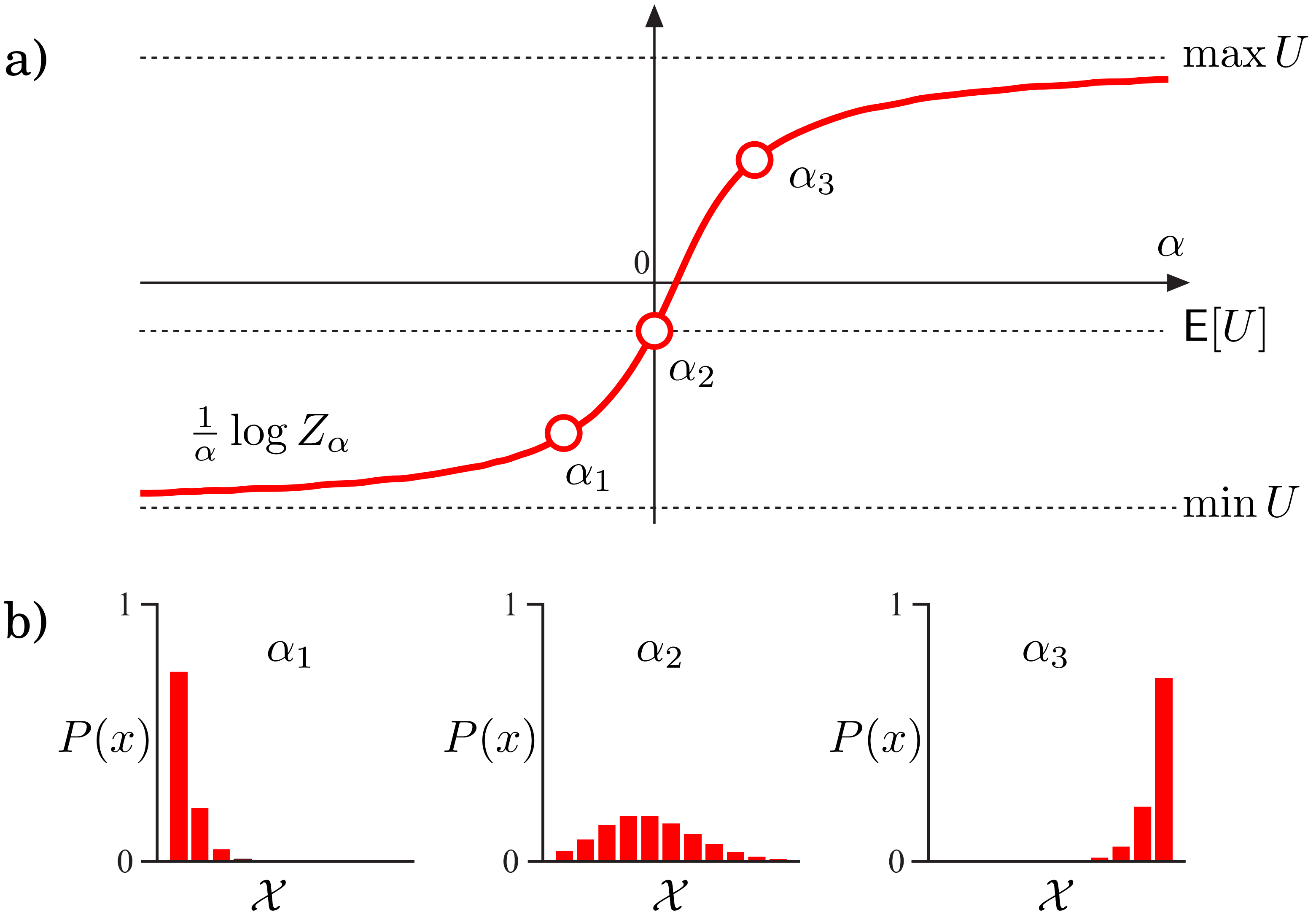
Optimal Choice and Certainty-Equivalent



Optimal Choice and Certainty-Equivalent

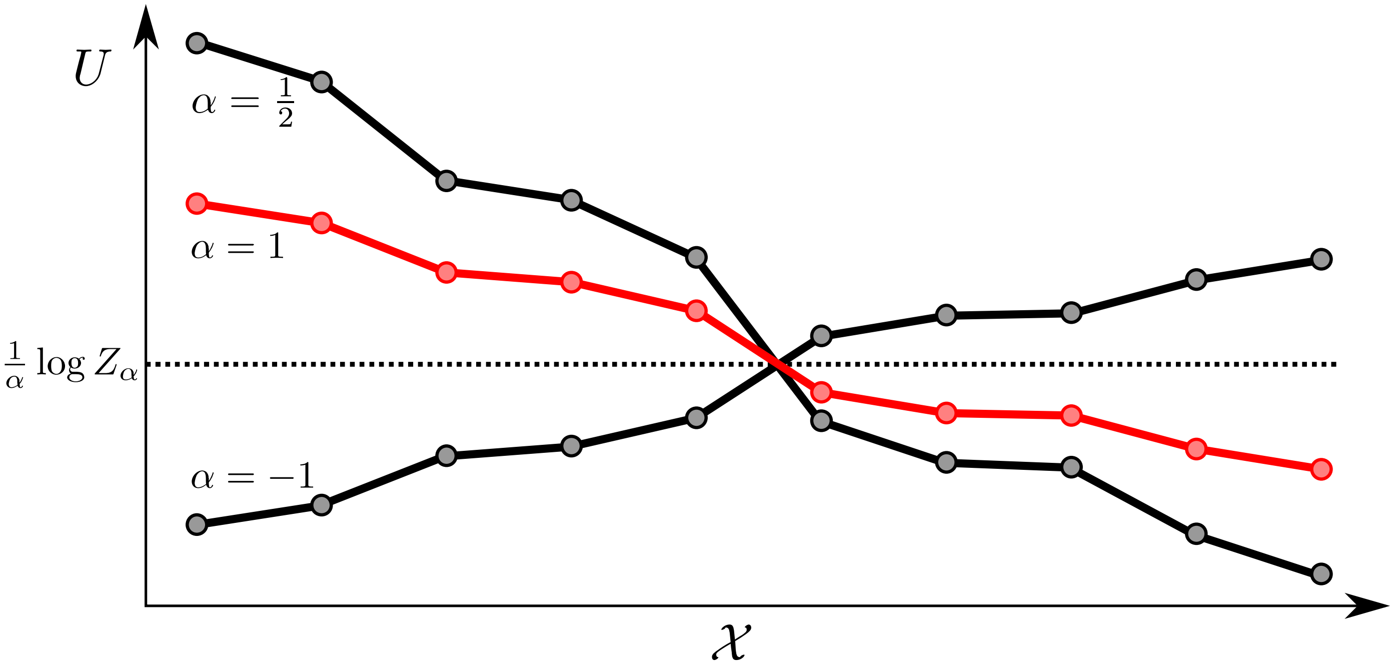


Optimal Choice and Certainty-Equivalent



Inverse temperature encapsulates constraints.

Equivalence



Behaviourally indistinguishable.

Have same:

- prior,
- posterior,
- and certainty-equivalent.

$$\alpha \longleftrightarrow \beta$$

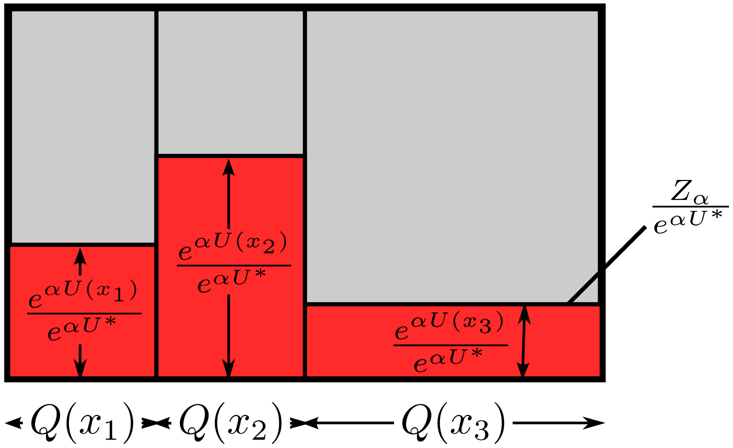
$$U \longleftrightarrow V$$

$$V(x) = \frac{\alpha}{\beta} U(x) + \left(\frac{1}{\alpha} - \frac{1}{\beta} \right) Z_\alpha$$

Many ways to represent same choice behaviour.

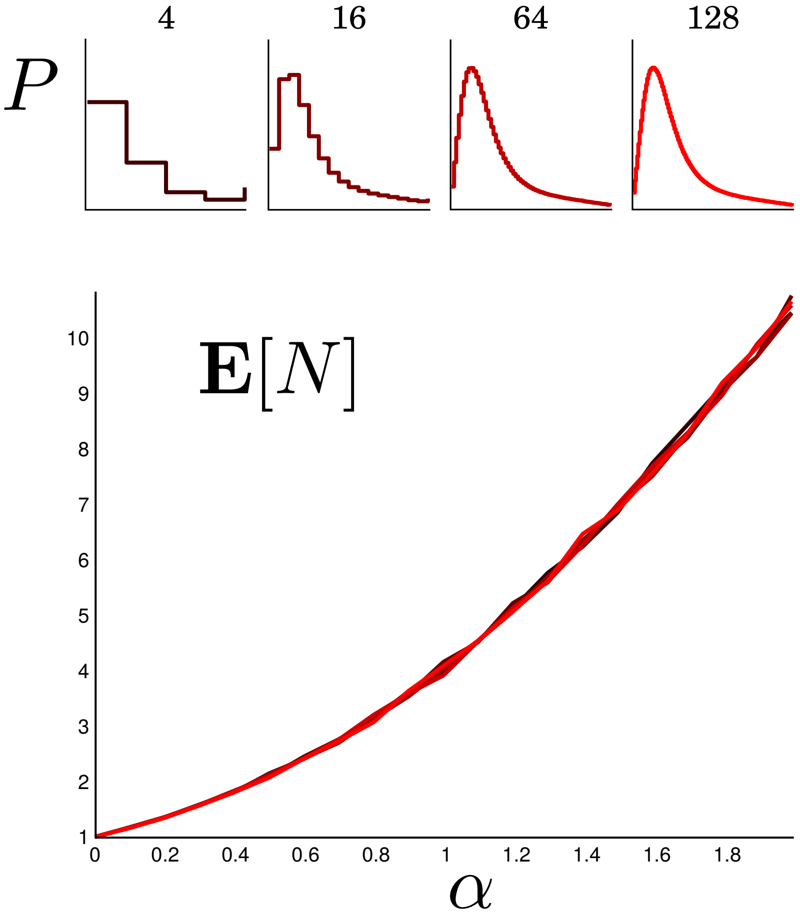
Stochastic Choices

Choosing amounts to **rejection sampling**:



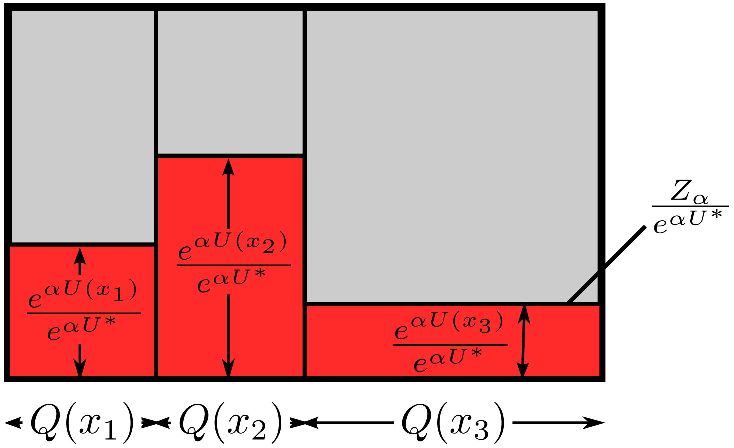
Input: U^*

Sample $u \sim \mathcal{U}(0,1)$
Sample $x \sim Q$
If $u \le e^{\alpha[U(x)-U^*]}$ accept,
else reject.



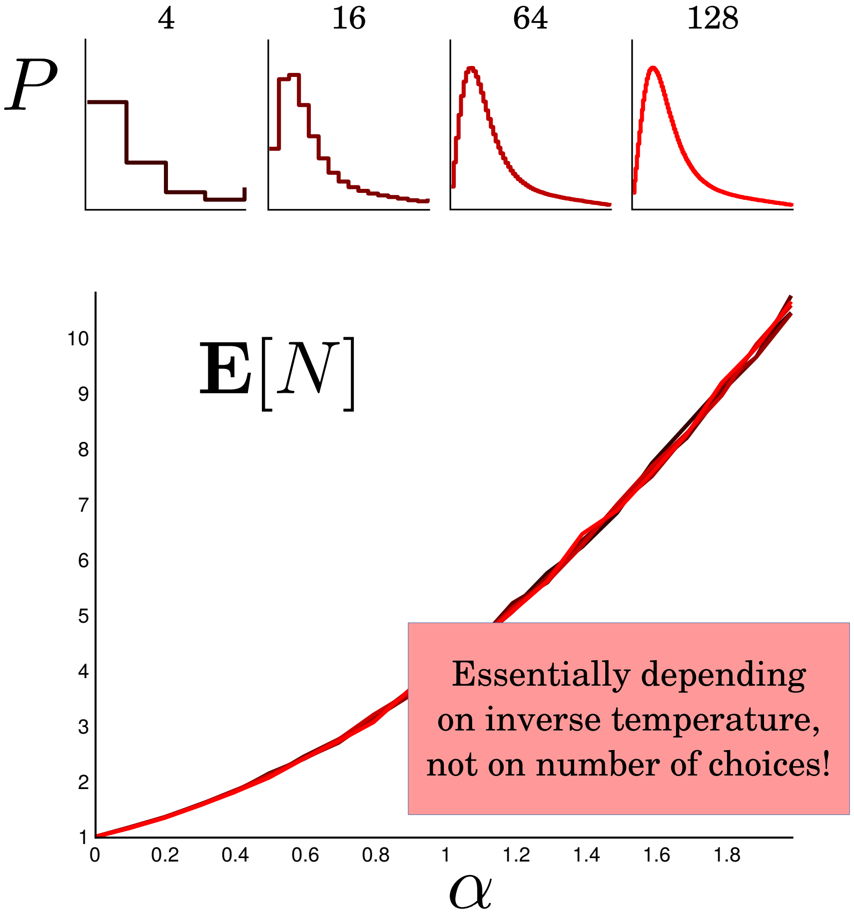
Stochastic Choices

Choosing amounts to **rejection sampling** [1]:



```
Input: $U^*$

Sample $u \sim \mathcal{U}(0,1)$
Sample $x \sim Q$
If $u \leq e^{\alpha[U(x)-U^*]}$ accept,
else reject.
```



Essentially depending on inverse temperature, not on number of choices!

RS is the most efficient sampler given knowledge.

Stochastic Choices II

We can sample directly from an equivalent decision problem.

$$p = \frac{Z_\alpha}{e^{\alpha U^*}} = \left(\frac{Z_\beta}{e^{\beta U^*}}\right)^{\frac{\alpha}{\beta}}$$

Interpretation

Obtain $\frac{\alpha}{\beta} \in \mathbb{R}$ samples from equivalent decision problem.

If **all** are accepted, then accept **any** as a sample of the original problem, otherwise reject.

Stochastic Choices II

We can sample directly from an equivalent decision problem.

$$p = \frac{Z_\alpha}{e^{\alpha U^*}} = \left(\frac{Z_\beta}{e^{\beta U^*}}\right)^{\frac{\alpha}{\beta}}$$

Interpretation

Obtain $\frac{\alpha}{\beta} \in \mathbb{R}$ samples from equivalent decision problem.

If **all** are accepted, then accept **any** as a sample of the original problem, otherwise reject.

RS from equivalent problem requires equalisation.

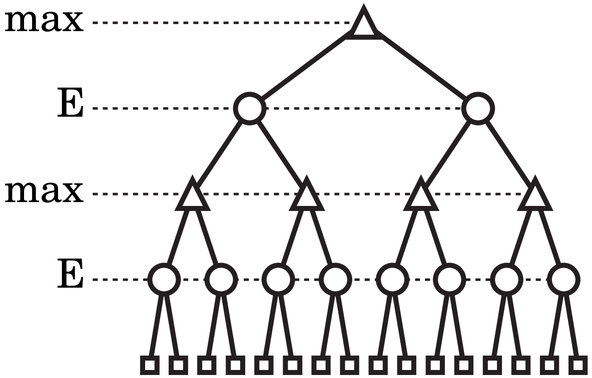
Comparison

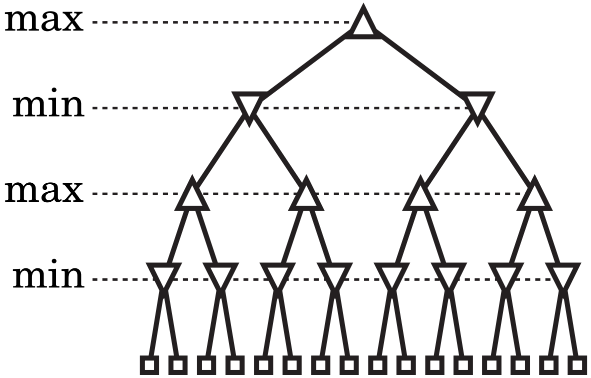
Objective Function	Expected Utility	Free Energy
Rationality Paradigm	perfect	bounded
Domain size	small	large
Search strategy	deterministic	randomised
Search scope	complete	incomplete
Functional form	linear	non-linear
Utility sensitivity	first moment	all moments

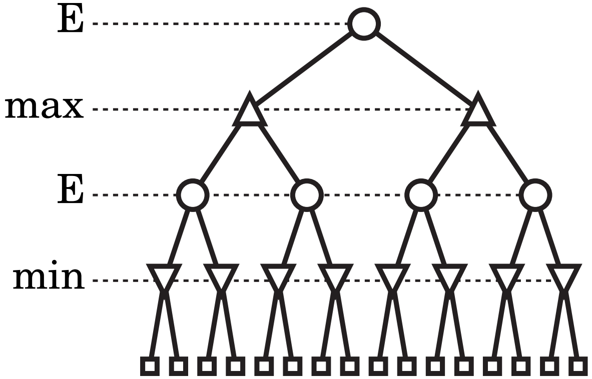
Sequential Decisions

Definition, construction, and recursions

Classical Decision Trees


Expectimax


Minimax

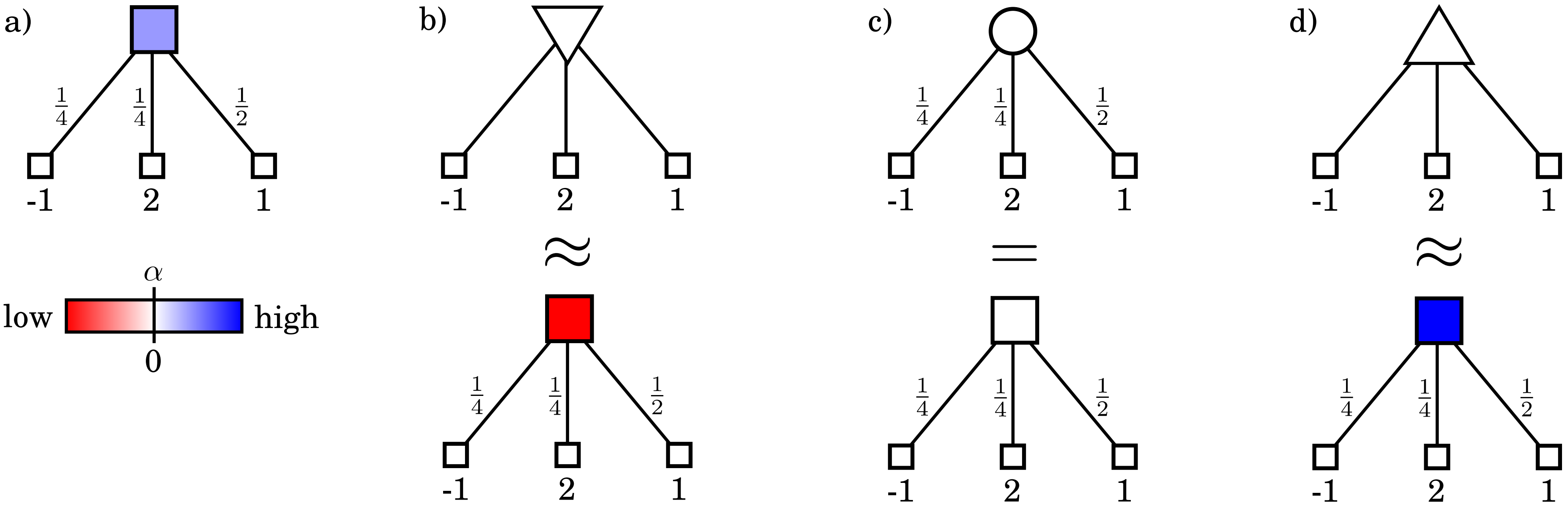

Expectiminimax

Planning in classical decision trees depends on the type of system making the move in each node.

Can we generalise them?

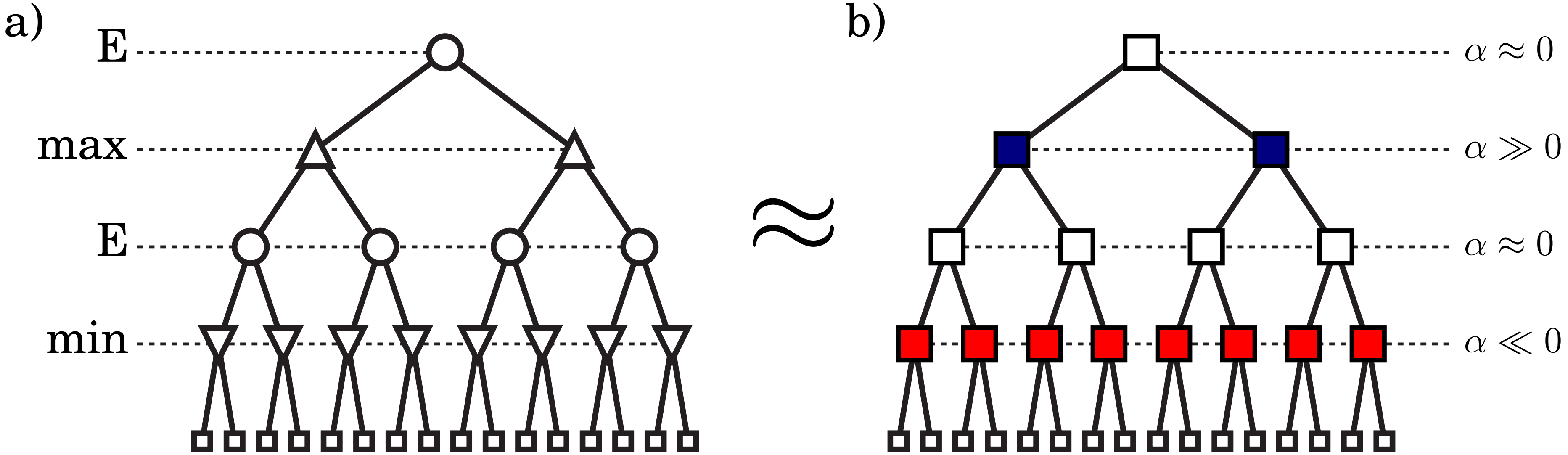
Certainty-Equivalent/Value

Classical values can be approximated



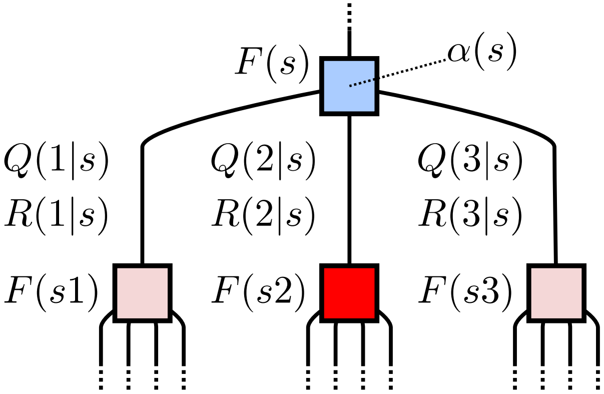
Different choices of inverse temperatures.

Bounded-Rational Decision Trees

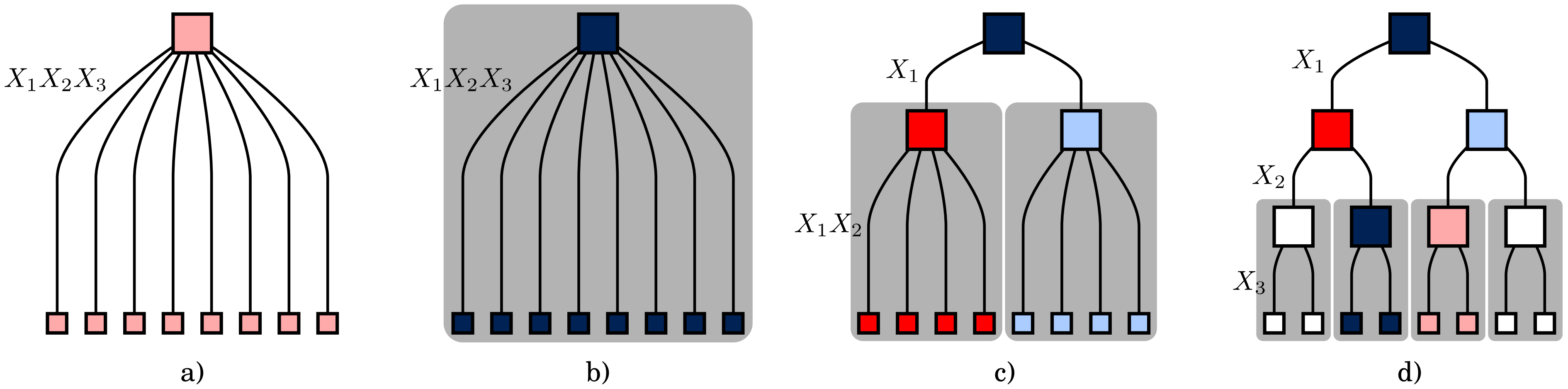


Free energy functional:

- **building block** for game trees
- inverse temperatures are **node-specific**
- solved using **forward sampling**, not dynamic programming



Construction of Decision Tree

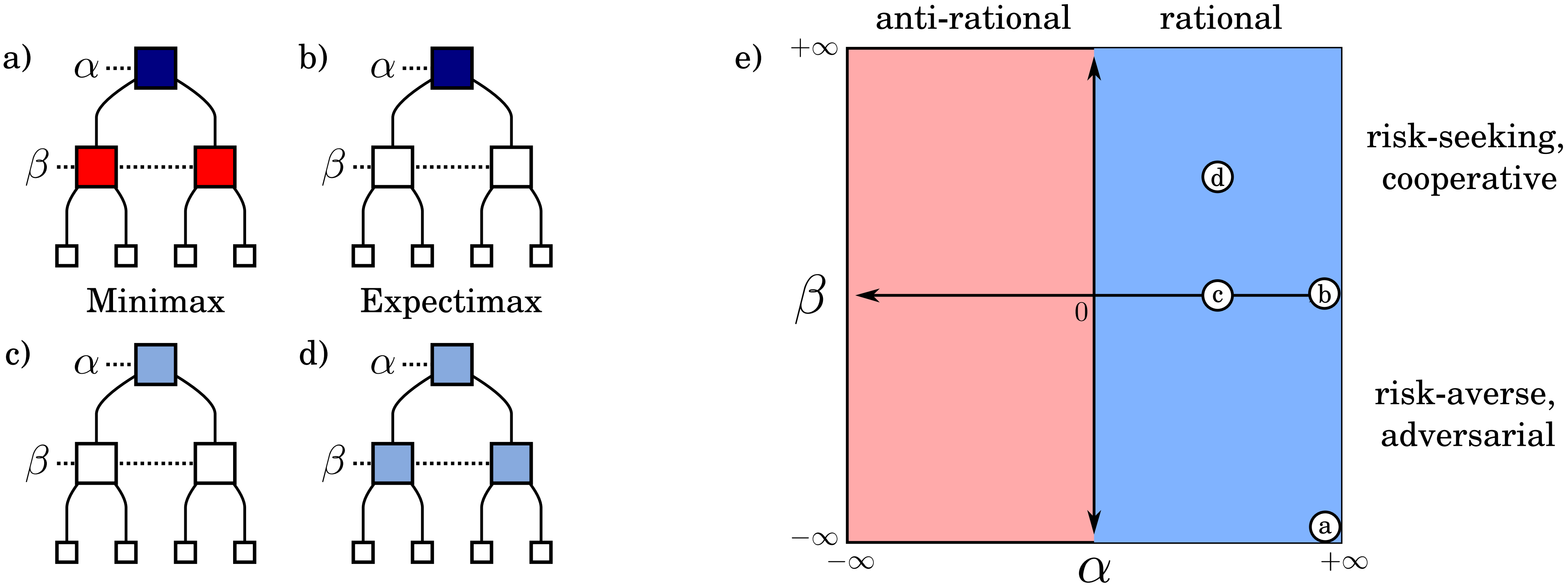


a) b) c) d)

Start from single-step.

Apply equivalence transformations until desired tree is obtained.

Map of Decision Rules



Inverse temperatures determine the decision rule.

Recursive Equations

Free Energy:

$$F[\tilde{P}](s) = \sum_x \tilde{P}(x|s)\left[R(x|s) - \frac{1}{\alpha(s)}\log\frac{\tilde{P}(x|s)}{Q(x|s)} + F[\tilde{P}](sx)\right]$$

Certainty-Equivalent:

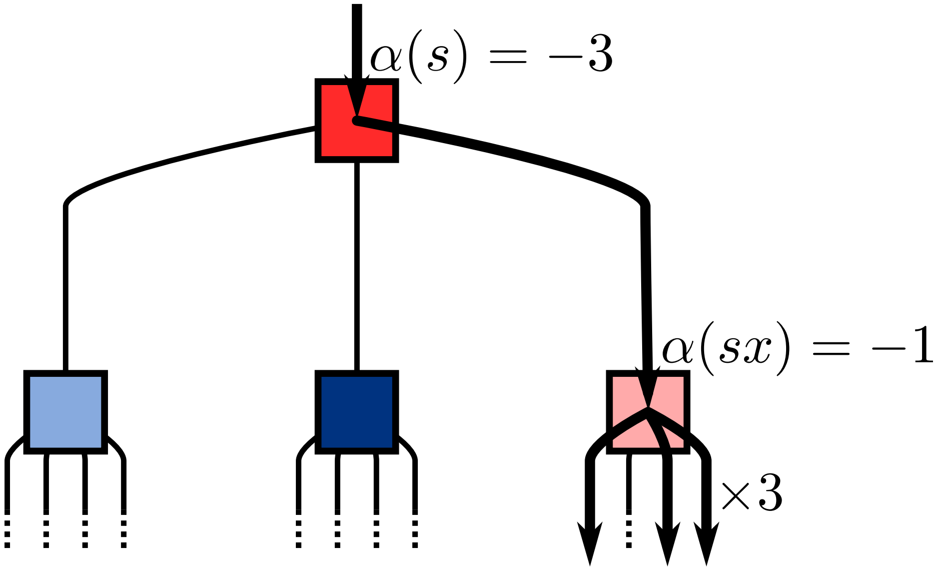
$$F(s) = \frac{1}{\alpha(s)}\log\sum_x Q(x|s)\exp\left\{\alpha(s)\left[R(x|s) + F(sx)\right]\right\}$$

Posterior:

$$P(x|s) = \frac{1}{Z(s)}Q(x|s)\exp\left\{\alpha(s)\left[R(x|s) + F(sx)\right]\right\}$$

Recursive Rejection Sampling

Recursively sample a path.
Equalise every time the inverse temperature changes.



```
**function** RRS$(s, U^*, \sigma)$

*Proposal:* $x \sim Q$

*Base case:*
If $sx$ is leaf then
    $u \sim \mathcal{U}(0,1)$
    If $(\sigma > 0$ and $u \leq p)$ then return $sx$
    If $(\sigma < 0$ and $1/u \leq p)$ then return $sx$
    Return $\epsilon$

*Recursion:*
$\sigma \leftarrow \sigma \cdot \mathrm{sign}(\alpha(s)/\alpha(sx))$
$\xi \leftarrow \mathrm{abs}(\alpha(s)/\alpha(sx))$
Generate $\xi$ samples from
    RRS$\big(sx, U^* - R(x|s), \sigma\big)$
If all successfull, then return any.
Return $\epsilon$
```

Implicitly evaluates feasibility of trajectories.

Summary

Problem:

- Large-scale decision spaces.

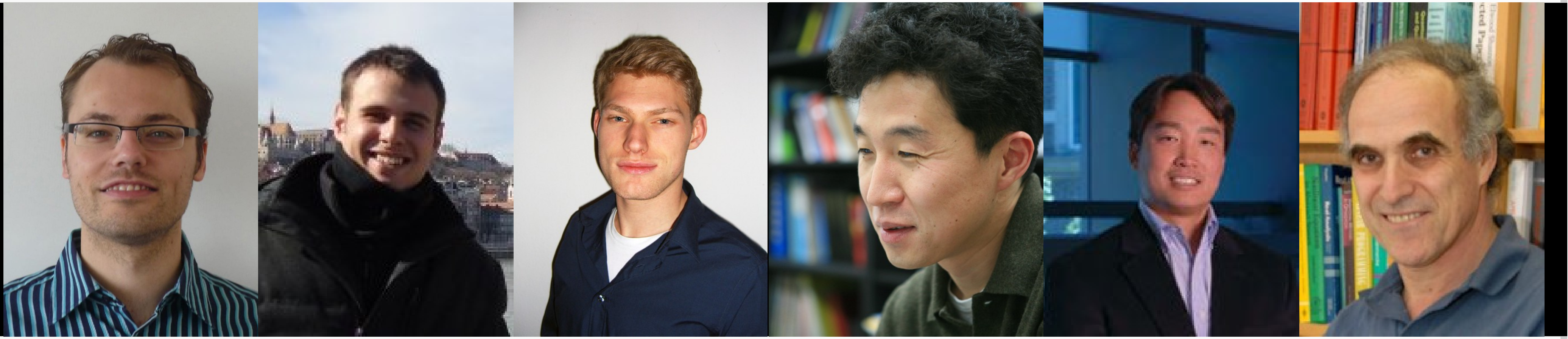
Key ideas:

- Use statistical description of choices.
- Exploit stochastic computational model.

Properties:

- Blurs boundary learning-planning.
- Massively parallelisable.
- Controllable complexity.
- Implements rich class of decision rules.

Thank you!



Bonus Material…

Meta-Reasoning

- Reason about costs of reasoning
- Penalise search complexity

$$
\begin{aligned}
&U(\pi) \\
U'(\pi, \pi') &:= U(\pi) - C(\pi, \pi')
\end{aligned}
$$

Meta-Reasoning

- Reason about costs of reasoning
- Penalise search complexity

$$\begin{aligned}
U(\pi) \\
U'(\pi, \pi') &:= U(\pi) - C(\pi, \pi') \\
U''(\pi, \pi', \pi'') &:= U'(\pi, \pi') - C(\pi, \pi', \pi'') \\
&\vdots
\end{aligned}$$

- Much harder problem!

Meta-Reasoning

- Reason about costs of reasoning
- Penalise search complexity

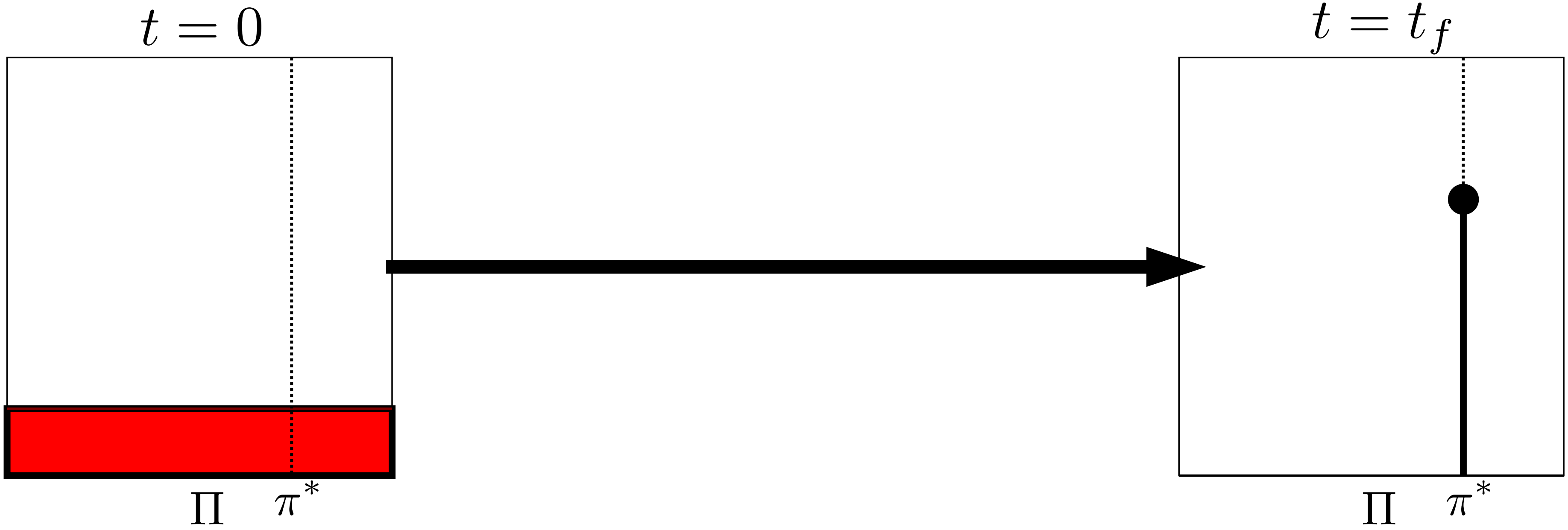
$$\begin{aligned} & U(\pi) \\ U'(\pi,\pi') &:= U(\pi) - C(\pi,\pi') \\ U''(\pi,\pi',\pi'') &:= U'(\pi,\pi') - C(\pi,\pi',\pi'') \\ & \vdots \end{aligned}$$

- Much harder problem!

Meta-reasoning is not permitted!

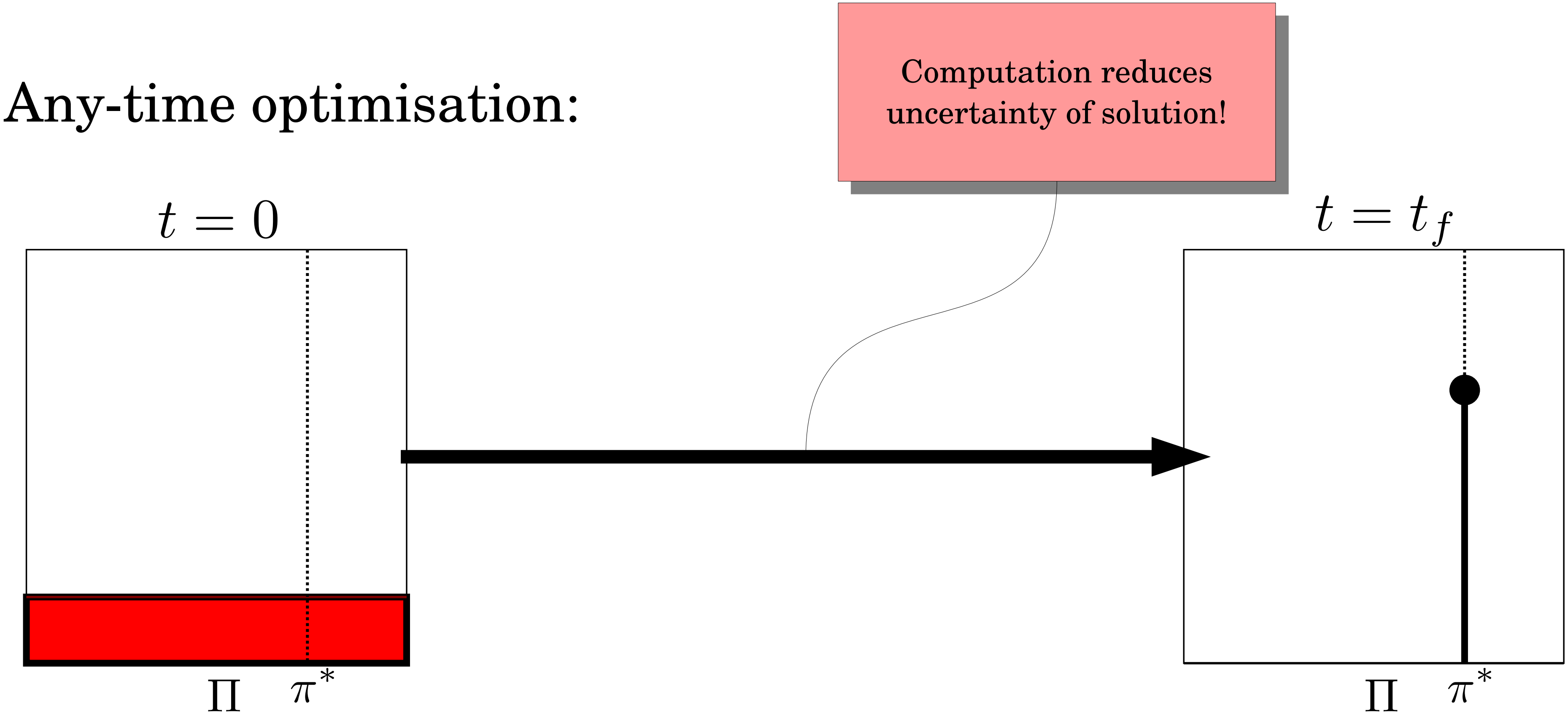
Interrupted Reasoning

Any-time optimisation:



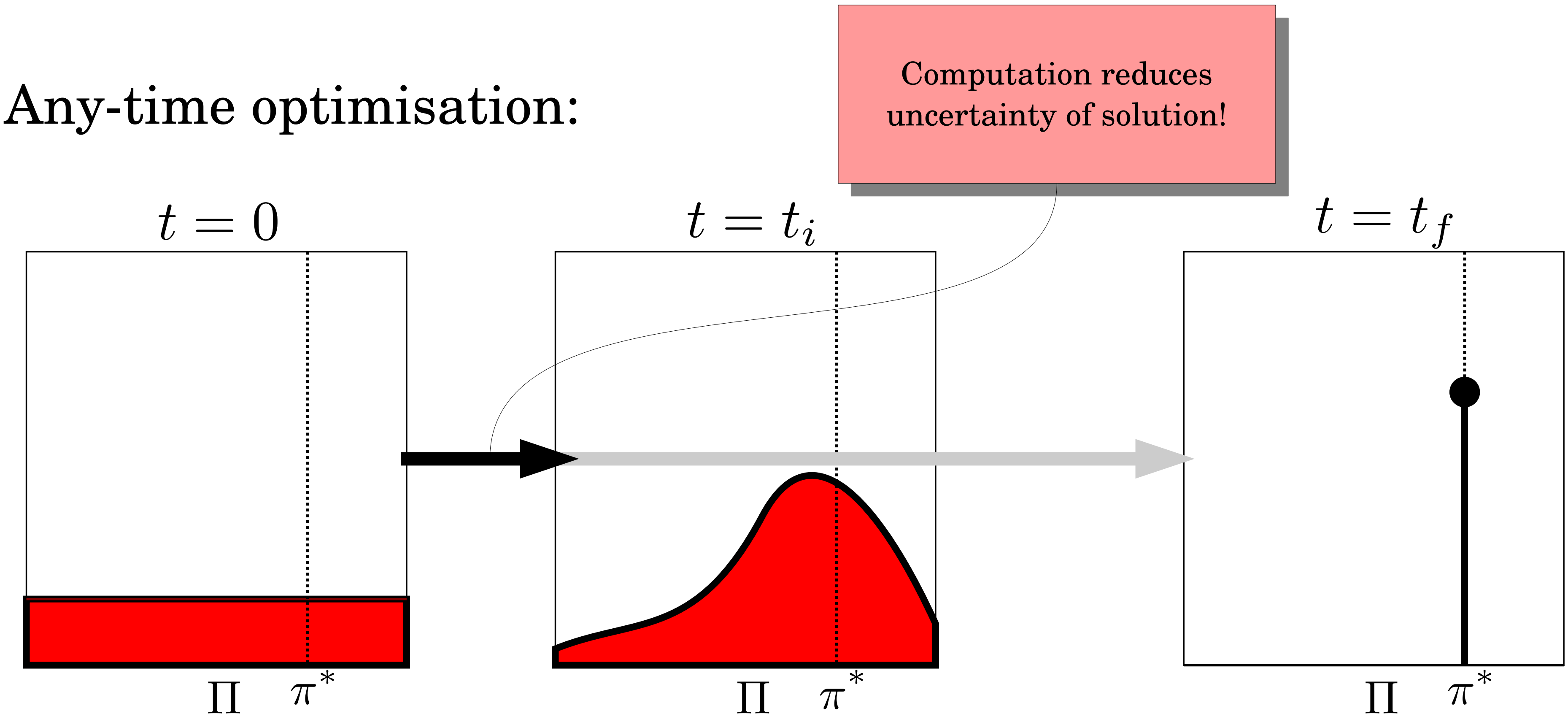
Interrupted Reasoning

Any-time optimisation:


Computation reduces uncertainty of solution!

Interrupted Reasoning

Any-time optimisation:



Computation reduces uncertainty of solution!

The Rational Adaptive Robobug

- Design:
  - Horizon: **120 seconds**
  - Frequency: **1 Hz**
  - Observations: **2**
  - Actions: **2**
- Size of decision tree:
  - Nodes: **10^{72}**
- Computation (@ 10^{16} FLOPS):
  - Time: **10^{50} years**

The Rational Adaptive Robobug

- Design:
  - Horizon: **120 seconds**
  - Frequency: **1 Hz**
  - Observations: **2**
  - Actions: **2**
- Size of decision tree:
  - Nodes: **10^{72}**
  - Atoms in the world: **10^{50}**
- Computation (@ 10^{16} FLOPS):
  - Time: **10^{50} years**
  - Age of the universe: **10^{10} years**

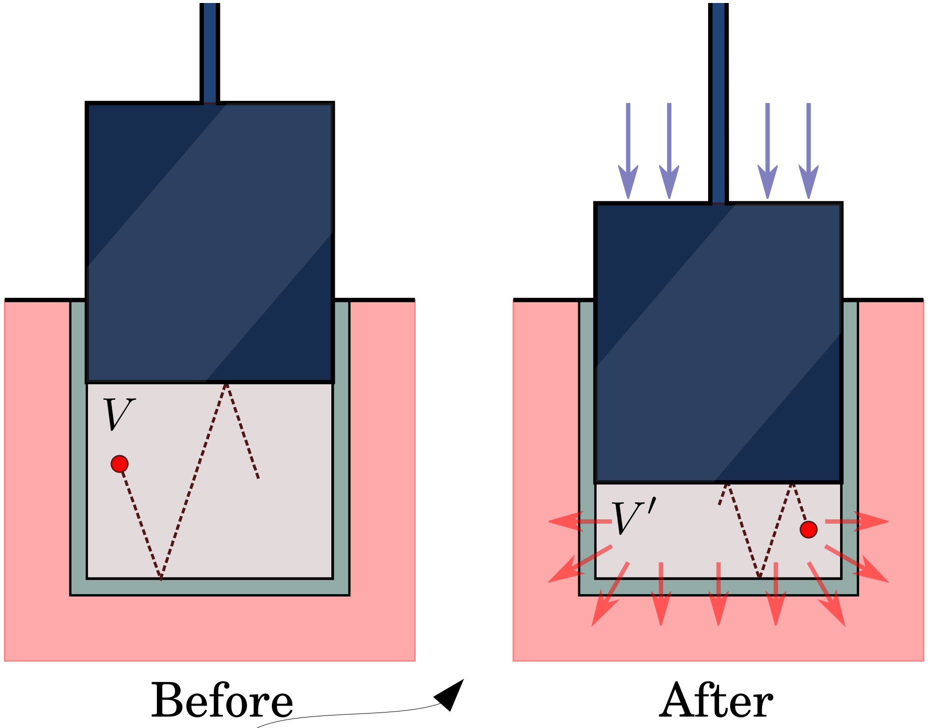
Relation Energy-Information

- Desiderata for costs
  - functions of probability
  - additive
  - monotonic
- Result [1]

$$C(E) = -\frac{1}{\beta}\log P(E)$$

Conversion Factor (real-valued)

Analogy: Isothermal Transformation



Work is reduction of uncertainty!

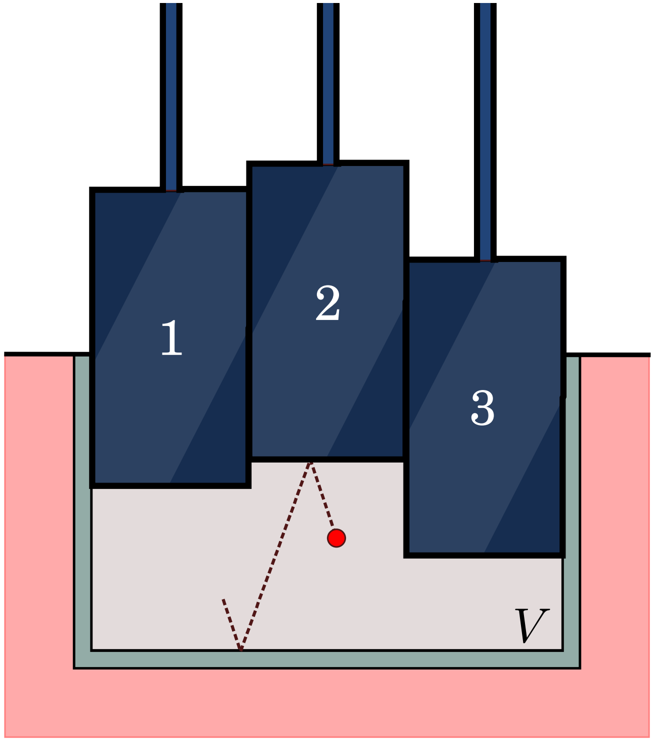
$$W = -\gamma \log \frac{V}{V'}$$

Utility = Work = (Unit Conv.) x (Reduction of Uncertainty).

[1] Ortega, P.A. & Braun, D.A., 2011

Energy-Information Trade-Off

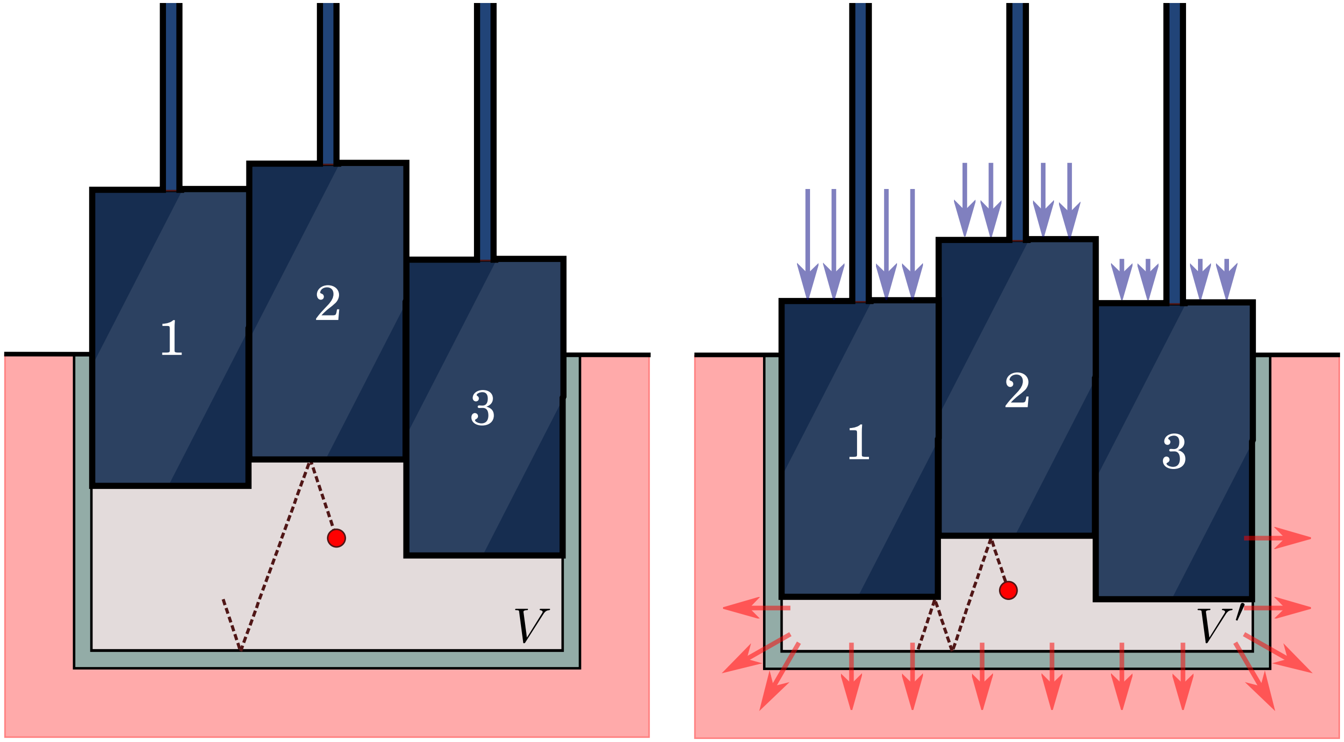
Analogy: Isothermal Transformation



Prior State

Energy-Information Trade-Off

Analogy: Isothermal Transformation



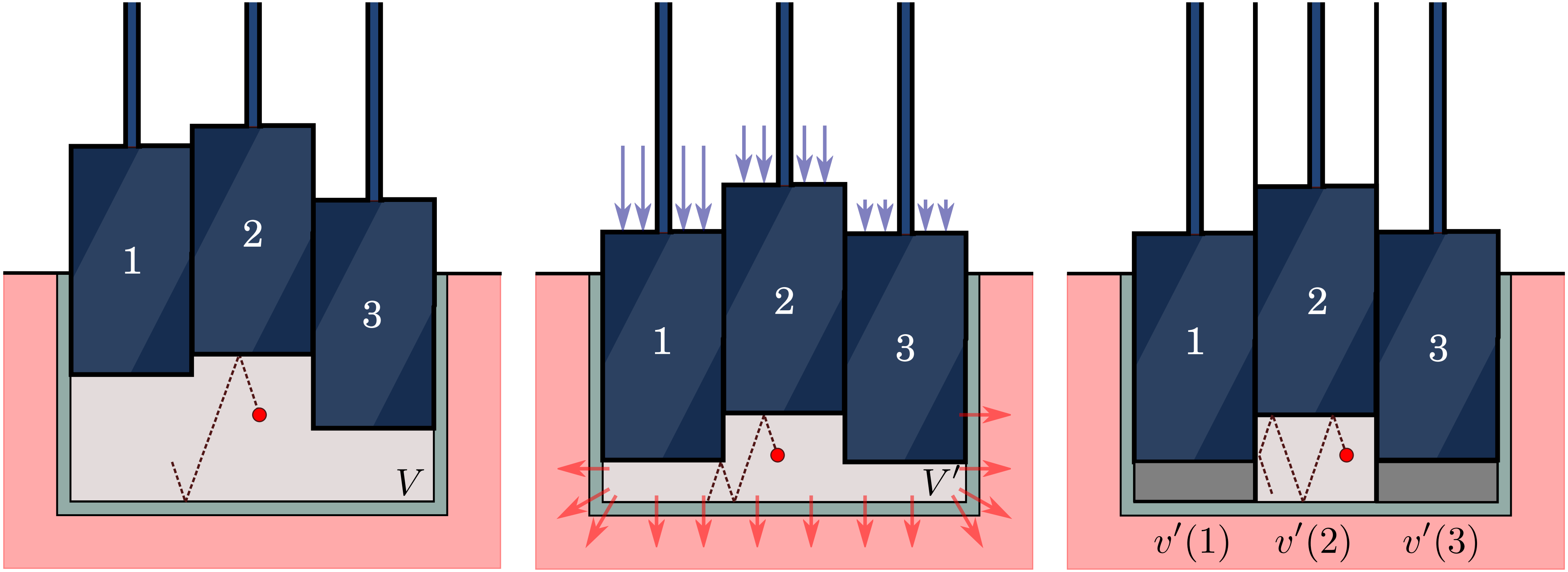
Prior State

Posterior State

$$W = -\gamma \log \frac{V}{V'}$$

Energy-Information Trade-Off

Analogy: Isothermal Transformation



Prior State

Posterior State

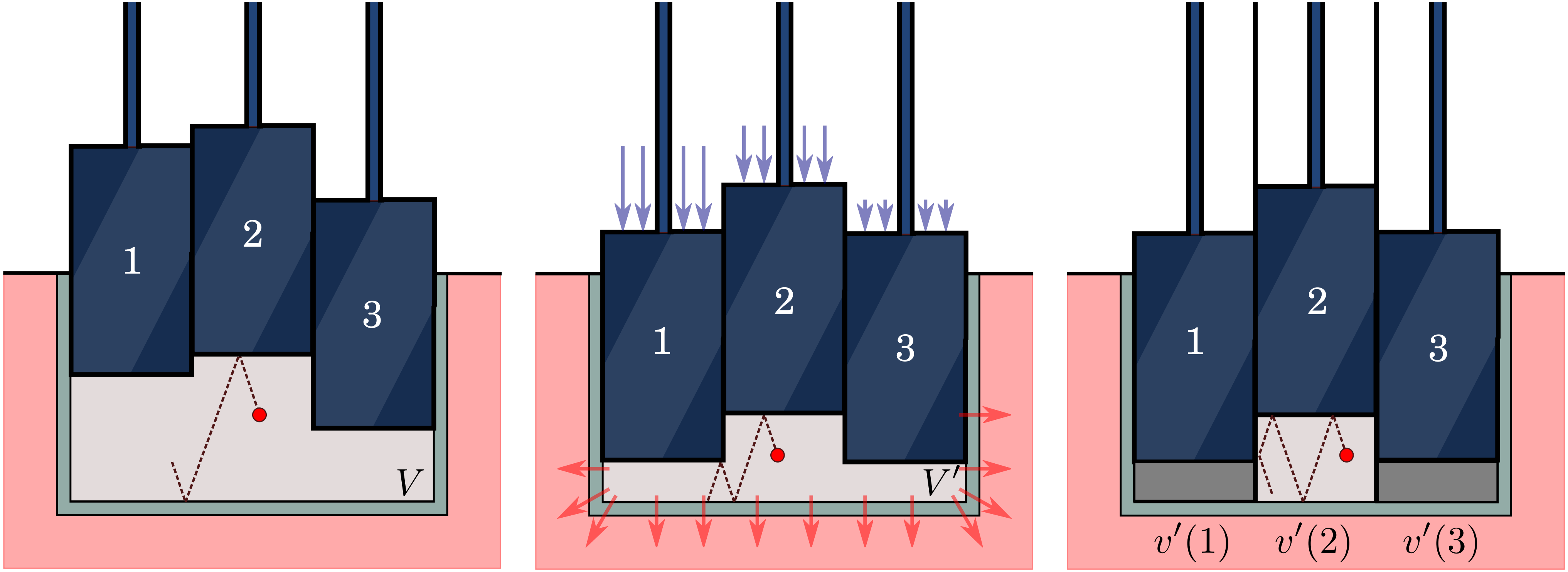
$$W = -\gamma \log \frac{V}{V'}$$

Realisation

$$w(2) = -\gamma \log \frac{v(2)}{v(2)'}$$

Energy-Information Trade-Off

Analogy: Isothermal Transformation



Prior State

Posterior State

$$W = -\gamma \log \frac{V}{V'}$$

Realisation

$$w(2) = -\gamma \log \frac{v(2)}{v(2)'}$$

Knowledge shapes the perceived work!

Knowledge shapes perceived work

Relation between work before/after knowledge:

$$W = -\gamma \log \frac{V}{V'}$$

Knowledge shapes perceived work

Relation between work before/after knowledge:

$$\begin{aligned} W &= -\gamma \log \frac{V}{V'} \\ &= -\gamma \sum_x \frac{v'(x)}{V'} \log \left\{ \frac{V}{V'} \cdot \frac{v'(x)}{v'(x)} \cdot \frac{v(x)}{v(x)} \right\} \\ &= -\gamma \sum_x p(x) \log \frac{v(x)}{v'(x)} - \gamma \sum_x p(x) \log \left\{ \frac{v'(x)}{V'} \cdot \frac{V}{v(x)} \right\} \\ &= \mathbb{E}_p[w] - \gamma \mathbb{D}(p \| q) \end{aligned}$$

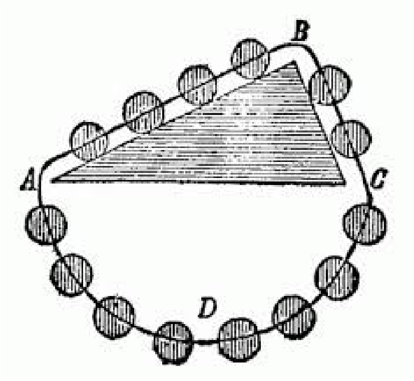
Knowledge shapes perceived work

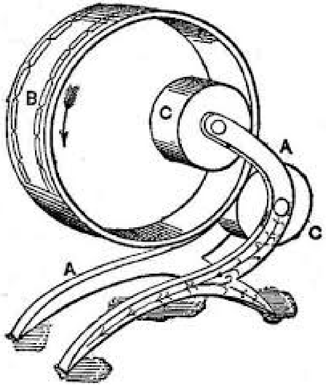
Relation between work before/after knowledge:

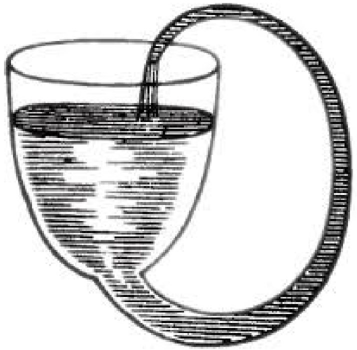
$$
\begin{aligned}
W &= -\gamma \log \frac{V}{V'} \\
&= -\gamma \sum_x \frac{v'(x)}{V'} \log\left\{ \frac{V}{V'} \cdot \frac{v'(x)}{v'(x)} \cdot \frac{v(x)}{v(x)} \right\} \\
&= -\gamma \sum_x p(x) \log \frac{v(x)}{v'(x)} - \gamma \sum_x p(x) \log\left\{ \frac{v'(x)}{V'} \cdot \frac{V}{v(x)} \right\} \\
&= \mathbb{E}_p[w] - \gamma \mathbb{D}(p \| q)
\end{aligned}
$$

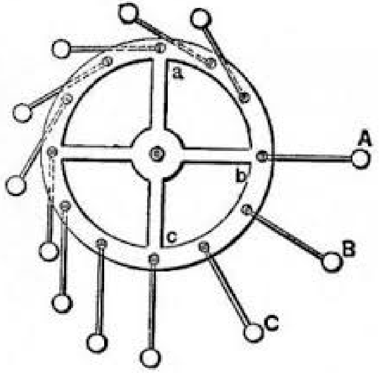
Free Energy = Work before knowledge of realisation.

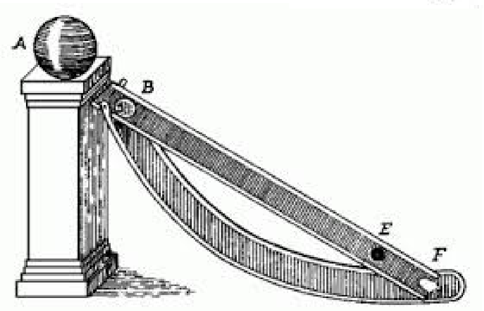
Lesson from Thermodynamics

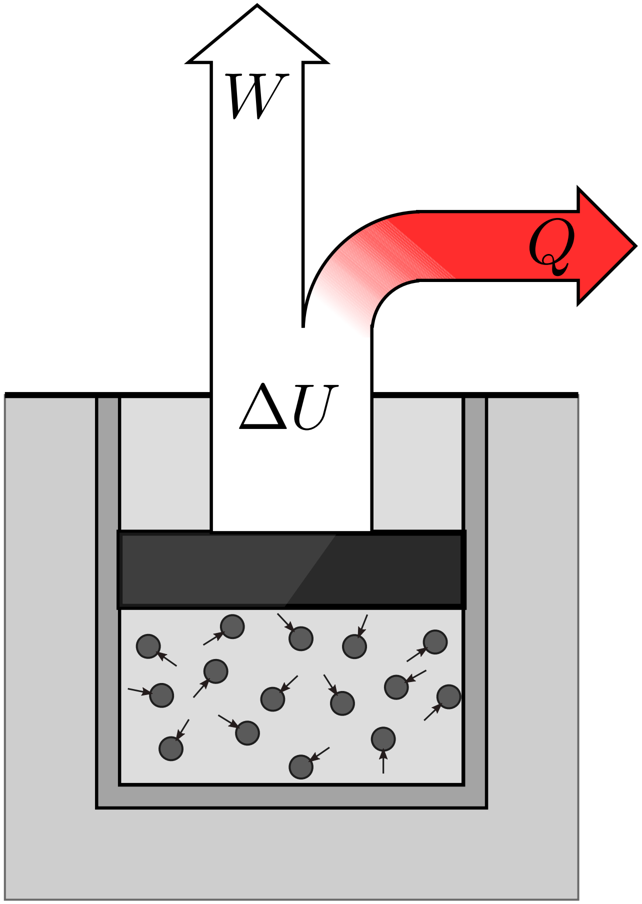












$$\Delta U = W + Q$$

No matter how hard we try, there's always a loss of energy.