Decision-Making with Information Constraints

Rate-Distortion and Hierarchical Decision-Making

Max Planck Institute for Intelligent Systems Tübingen Campus

Graduate Training Centre of Neuroscience
International Max Planck Research School

Eberhard Karls Universität Tübingen

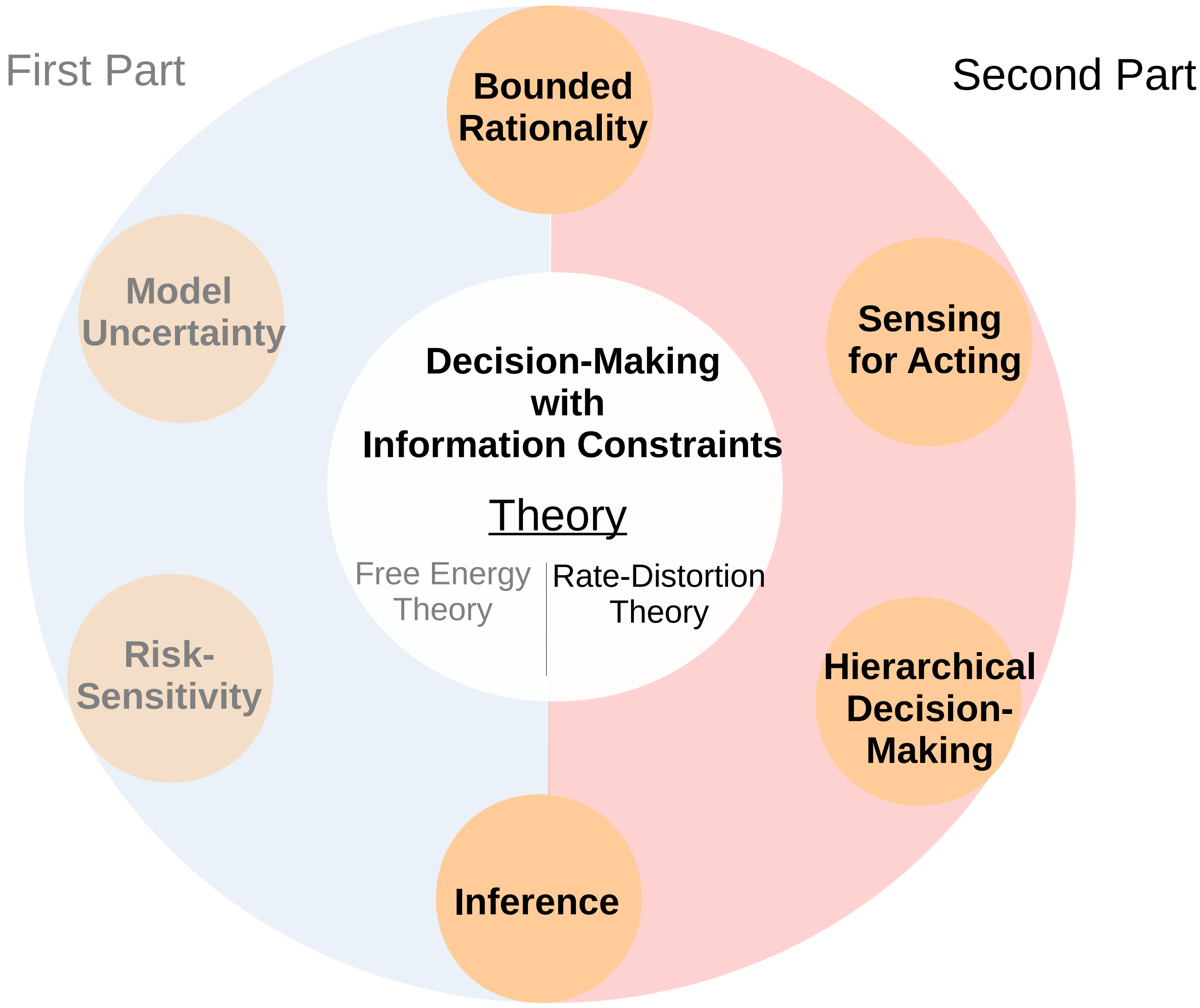
Max Planck Institute for Biological Cybernetics

Jordi Grau-Moya
PhD Student

Tim Genewein
PhD Student

June 2016, Philadelphia

Outline


First Part
Second Part
Bounded Rationality
Model Uncertainty
Sensing for Acting
Decision-Making with Information Constraints
Theory
Free Energy Theory
Rate-Distortion Theory
Risk-Sensitivity
Hierarchical Decision-Making
Inference

Quick recap: free energy principle for bounded rational decision-making

The problem with utility maximization

Goal:

- Given some world state w, pick best action a
- Desirability of action is specified by utility function $U(w, a)$

Easy...

$$a_w^* = \arg\max_a U(w, a)$$

$$w \longrightarrow a$$

The problem with utility maximization

Goal:

- Given some world state w, pick best action a
- Desirability of action is specified by utility function $U(w, a)$

Easy?

$$a_w^* = \arg\max_a U(w, a)$$

Problem:

- Searching through a vast set with limited computational capacity
- Finding the best action can easily become intractable

Bounded rational decision-making

Goal:

- Given some world state w, pick best action a
- Desirability of action is specified by utility function $U(w, a)$

Modify the optimality principle

- Take the process of computation into account
- Information-theoretic bounded rationality
- Rather than finding the single best action, find “good” actions that are actually computable

Information-theoretic bounded rationality

Find a stochastic policy $p(a|w)$ that maximizes

- expected utility $\sum_a p(a|w)U(w,a)$

subject to the constraint:

- “computational effort” $\leq K$

Information-theoretic bounded rationality

Find a stochastic policy $p(a|w)$ that maximizes

- expected utility $\sum_a p(a|w)U(w,a)$

subject to the constraint:

- "computational effort" $\leq K$

Computational effort?

- Transformation of behavior in response to observation
- Any change of behavior requires computation

$$\text{computational effort} = D_{\mathrm{KL}}(p(a|w)||p_0(a))$$

Information-theoretic bounded rationality

Trade off:

- Large expected utility
- Low computational effort

$$\boxed{\mathrm{p}^*(\mathrm{a}|\mathrm{w}) = \underset{p(a|w)}{\arg\max}\, \mathbf{E}_{p(a|w)}[U(w,a)] - \frac{1}{\beta} D_{\mathrm{KL}}(p(a|w)||p_0(a))}$$

Solving the variational problem

Trade off:

- Large expected utility
- Low computational effort

$$\mathrm{p}^*(\mathrm{a}|\mathrm{w}) = \underset{p(a|w)}{\arg\max}\, \mathbf{E}_{p(a|w)}[U(w,a)] - \frac{1}{\beta} D_{\mathrm{KL}}(p(a|w)||p_0(a))$$

$$\boxed{p^*(a|w) = \frac{1}{Z} p_0(a) e^{\beta U(w,a)}}$$

$Z(w)$ … partition sum, acts as normalization constant $\sum_a p_0(a) e^{\beta U(w,a)}$

β … inverse temperature, governs trade-off

Solving the variational problem

Trade off:

- Large expected utility
- Low computational effort

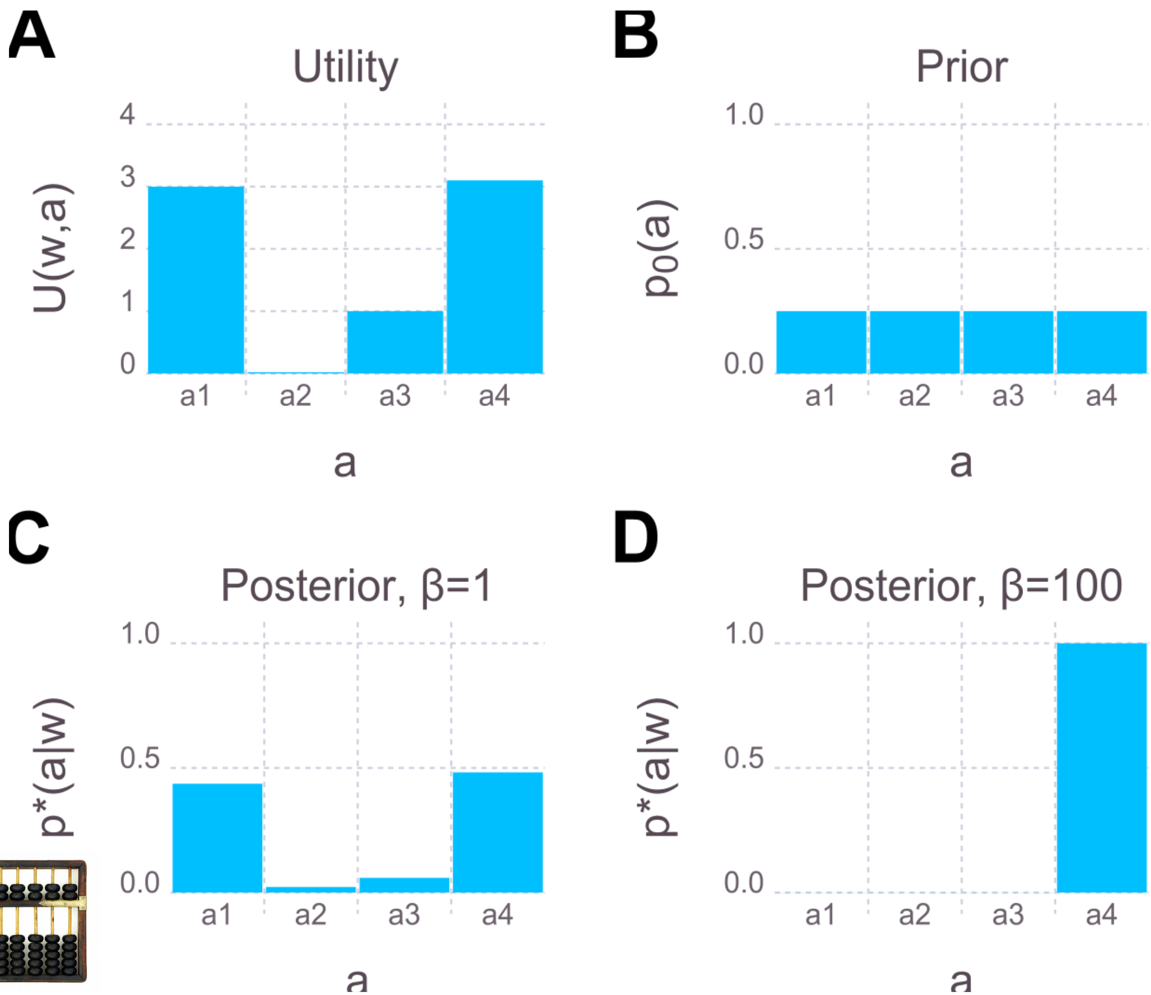
$$\mathrm{p}^*(\mathrm{a}|\mathrm{w}) = \underset{p(a|w)}{\arg\max}\, \mathbf{E}_{p(a|w)}[U(w,a)] - \frac{1}{\beta} D_{\mathrm{KL}}(p(a|w)||p_0(a))$$

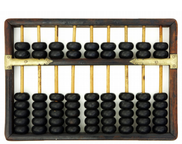
$$p^*(a|w) = \frac{1}{Z} p_0(a) e^{\beta U(w,a)}$$

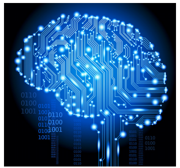
Special case: Bayes' rule: $U(w,a) = \log q(w|a)\,,\qquad \beta = 1$

$$p^*(a|w) = \frac{q(w|a)p_0(a)}{Z}$$

Example: grasping movement

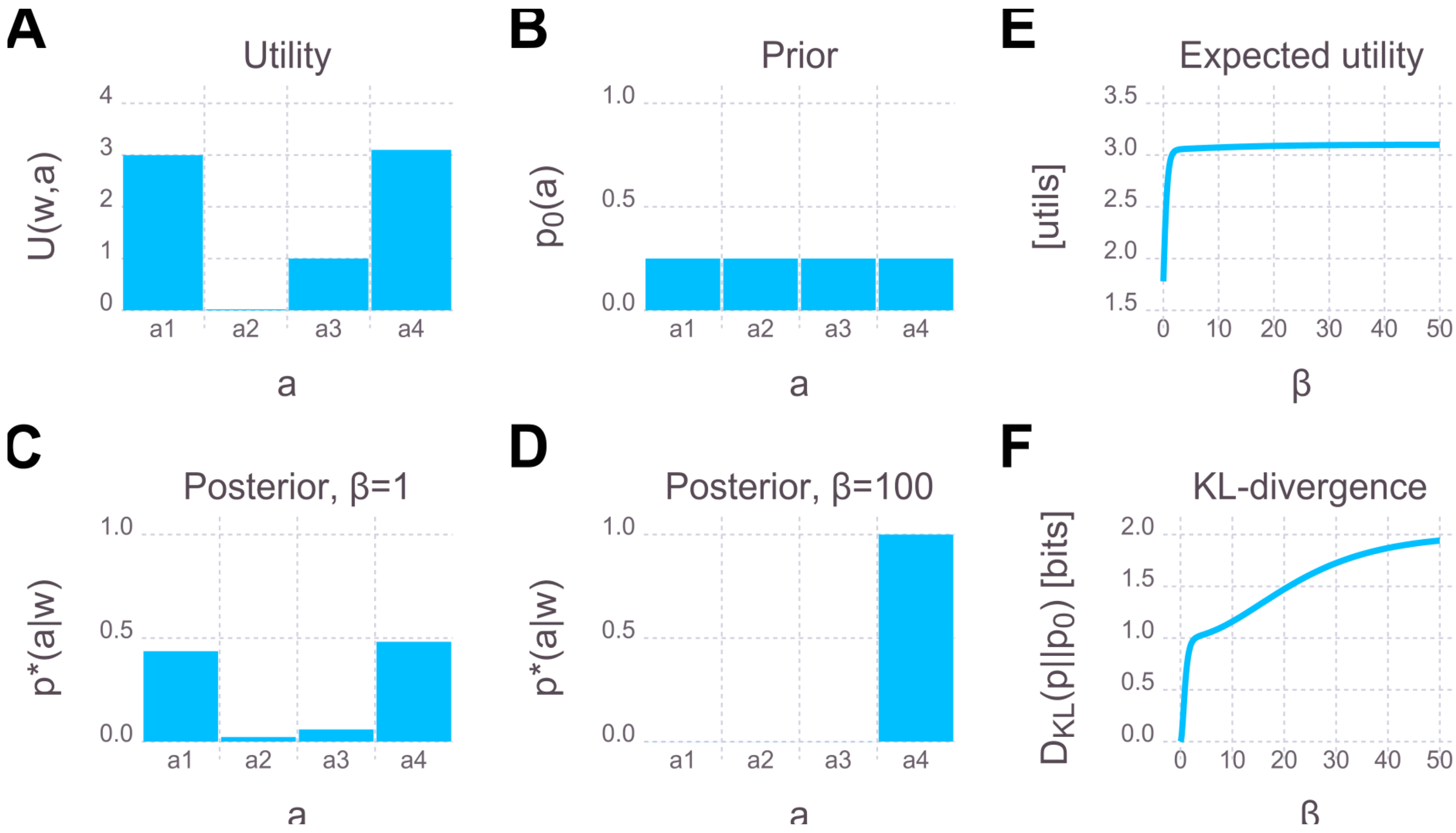






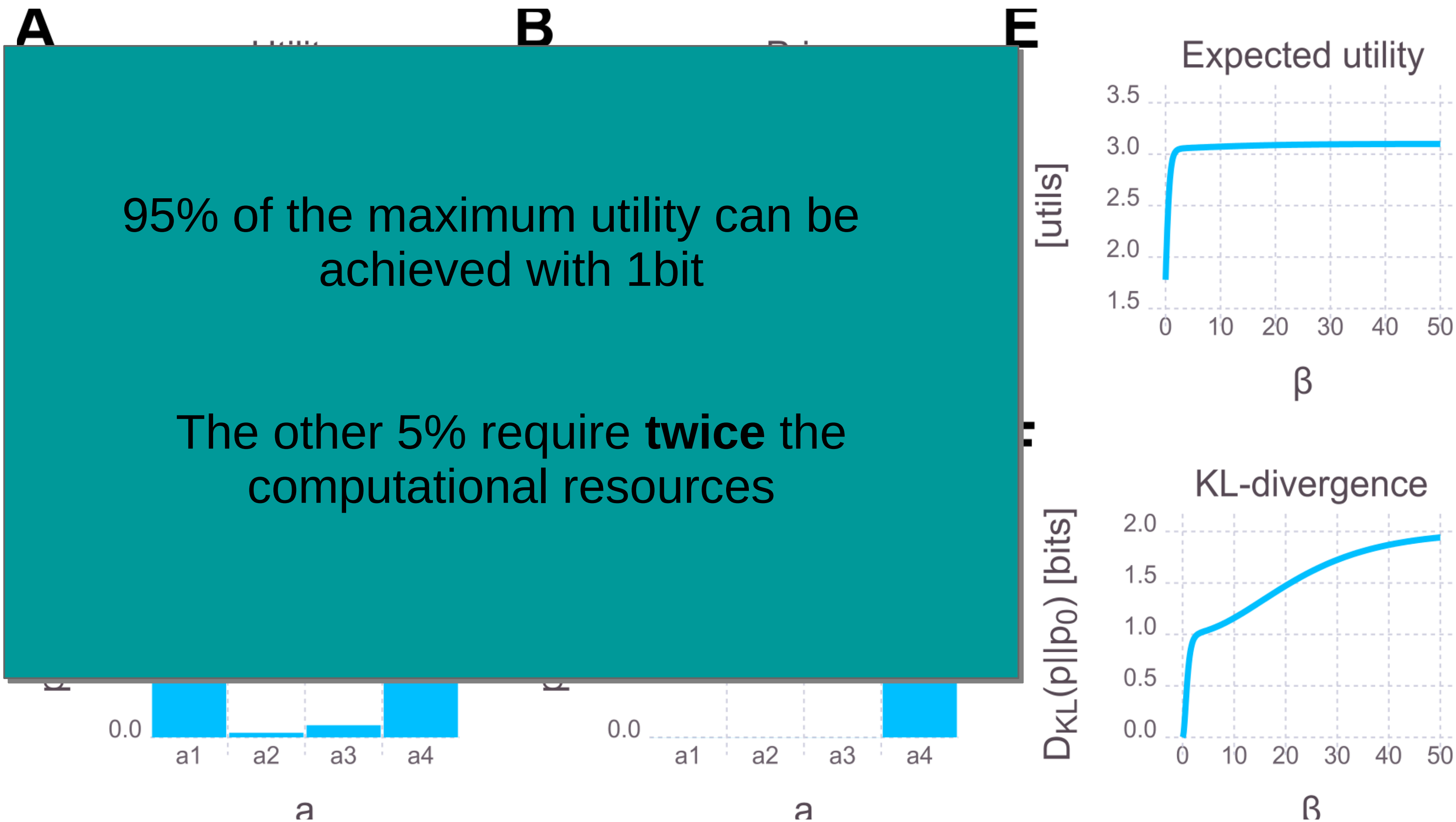
Example: grasping movement

Inverse temperature β governs computational resources



Example: grasping movement

Inverse temperature β governs computational resources



95% of the maximum utility can be achieved with 1bit

The other 5% require **twice** the computational resources

Rate-distortion theory for decision-making

Multiple world-states

Trade off large utility against low computational effort

- Now: consider multiple w, more precisely $p(w)$
- What is the optimal prior $p_0(a)$
- What is the (bounded) optimal $p^*(a|w)$

Reminder: free-energy principle

$$\mathrm{p}^*(\mathrm{a}|\mathrm{w}) = \underset{p(a|w)}{\arg\max}\, \mathbf{E}_{p(a|w)}[U(w,a)] - \frac{1}{\beta} D_{\mathrm{KL}}(p(a|w)||p_0(a))$$

Multiple world-states

Trade off large utility against low computational effort

- Now: consider multiple w, more precisely $p(w)$
- What is the optimal prior $p_0(a)$
- What is the (bounded) optimal $p^*(a|w)$

$$p^*(a|w) = \underset{p(a|w)}{\arg\max}\, \mathbf{E}_{p(w,a)}[U(w,a)] - \frac{1}{\beta} I(W;A)$$

- Computational effort = mutual information
  - Optimal prior is marginal $p(a) = \sum_w p(w)p(a|w)$
  - Mutual information is average KL-divergence

Multiple world-states

Trade off large utility against low computational effort

- Now: consider multiple w, more precisely $p(w)$
- What is the optimal prior $p_0(a)$
- What is the (bounded) optimal $p^*(a|w)$

$$p^*(a|w) = \arg\max_{p(a|w)} \mathbf{E}_{p(w,a)}[U(w,a)] - \frac{1}{\beta} I(W;A)$$

- Mathematically equivalent to rate-distortion problem
  - Framework for lossy compression
  - Communication-channel from world-states to actions with limited capacity

Multiple world-states

Trade off large utility against low computational effort

- Now: consider multiple w, more precisely $p(w)$
- What is the optimal prior $p_0(a)$
- What is the (bounded) optimal $p^*(a|w)$

$$p^*(a|w) = \underset{p(a|w)}{\arg\max}\, \mathbf{E}_{p(w,a)}[U(w,a)] - \frac{1}{\beta} I(W;A)$$

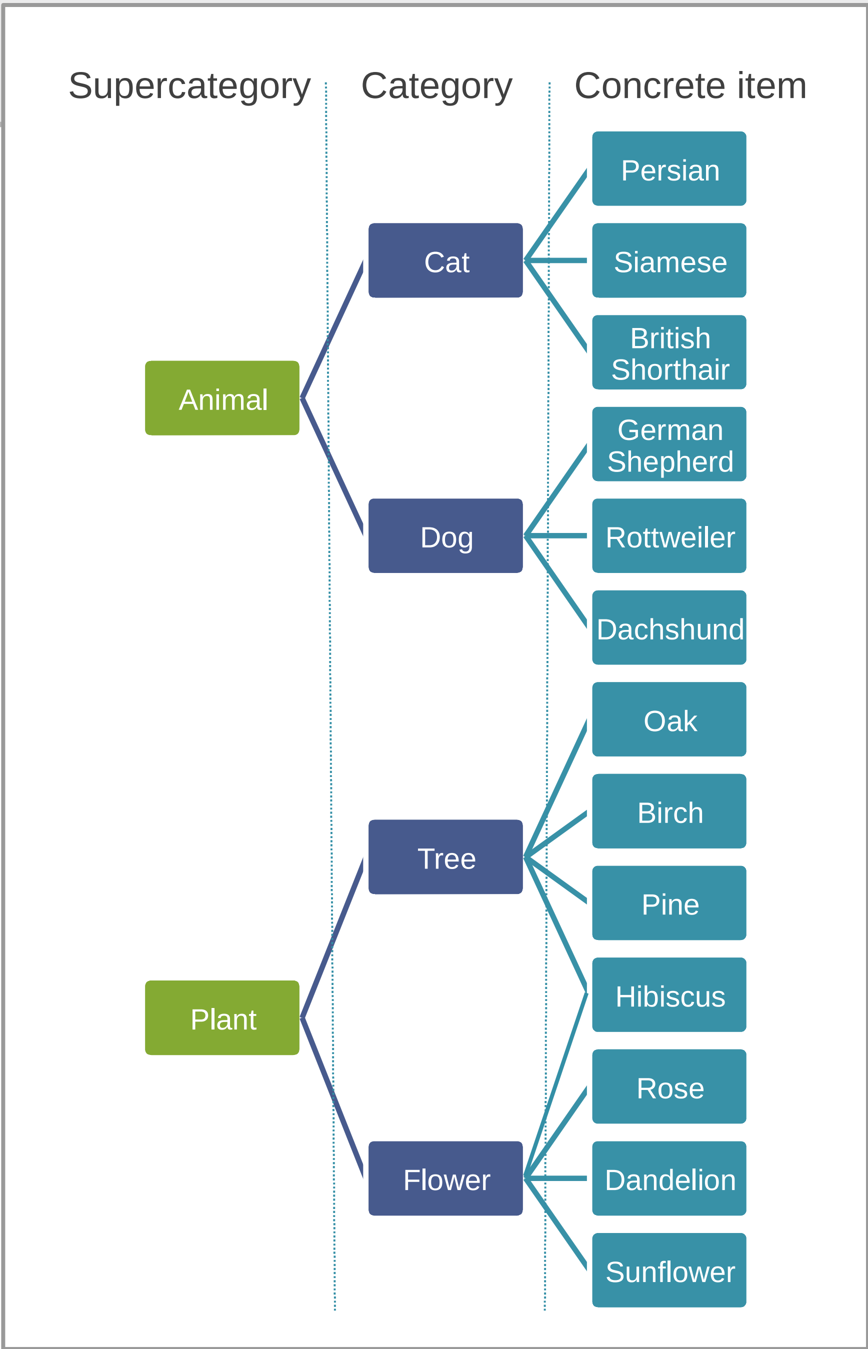
Solution:

$$\boxed{\begin{aligned} p^*(a|w) &= \frac{1}{Z} p(a) e^{\beta U(w,a)} \\ p(a) &= \sum_w p(w) p^*(a|w) \end{aligned}}$$

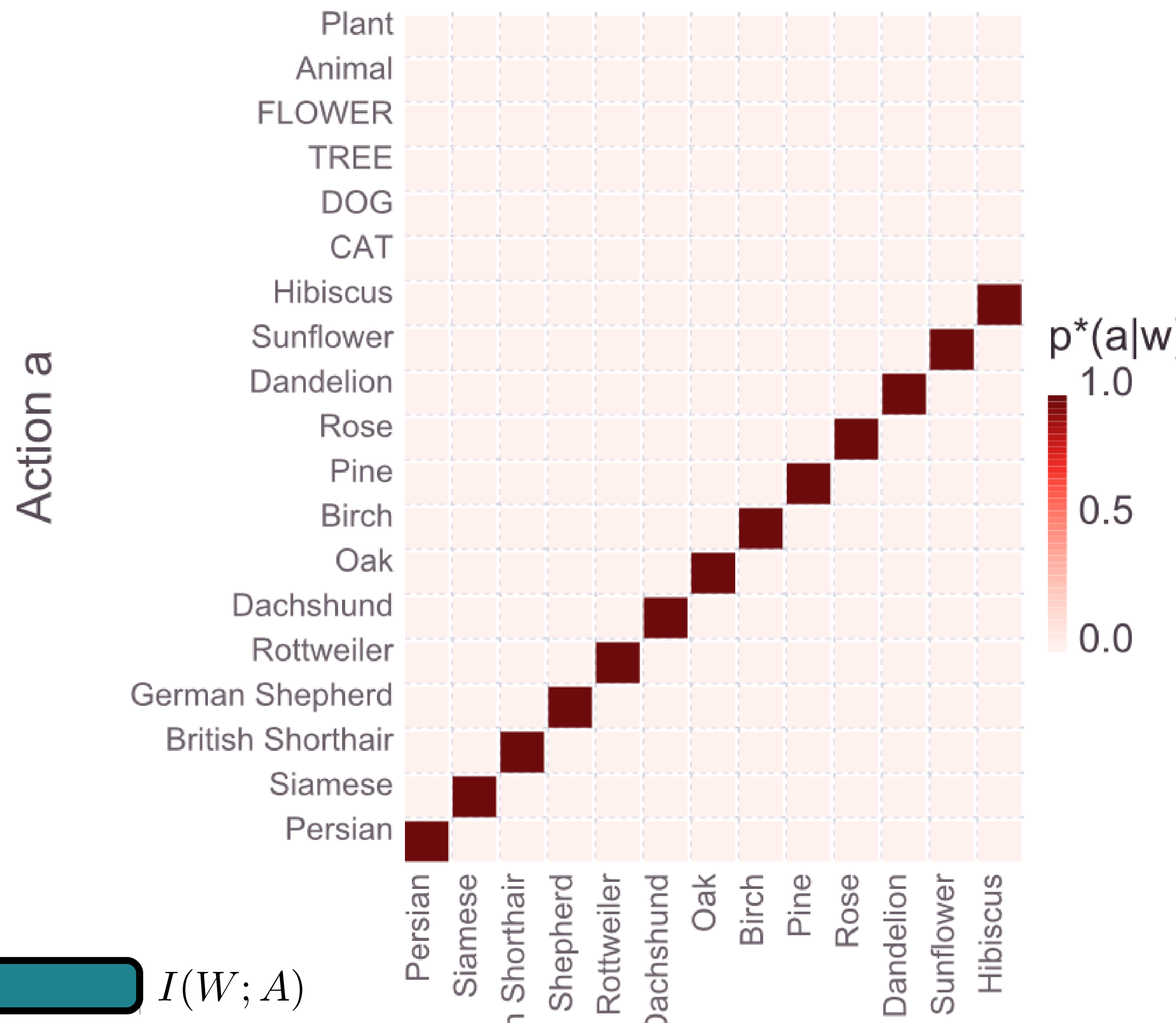
$$Z = \sum_a p(a) e^{\beta U(w,a)}$$

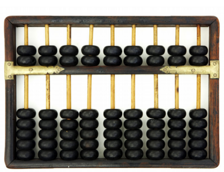
Toy Example

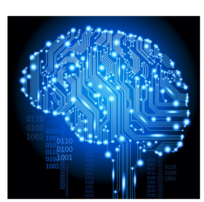
- Sensory state w:
  - Concrete item
- Action
  - Concrete item
  - Category
  - Supercategory
- Rewards/Utilities:
  - 3 if concrete item correct
  - 2.2 if category correct
  - 1.6 if supercategory correct



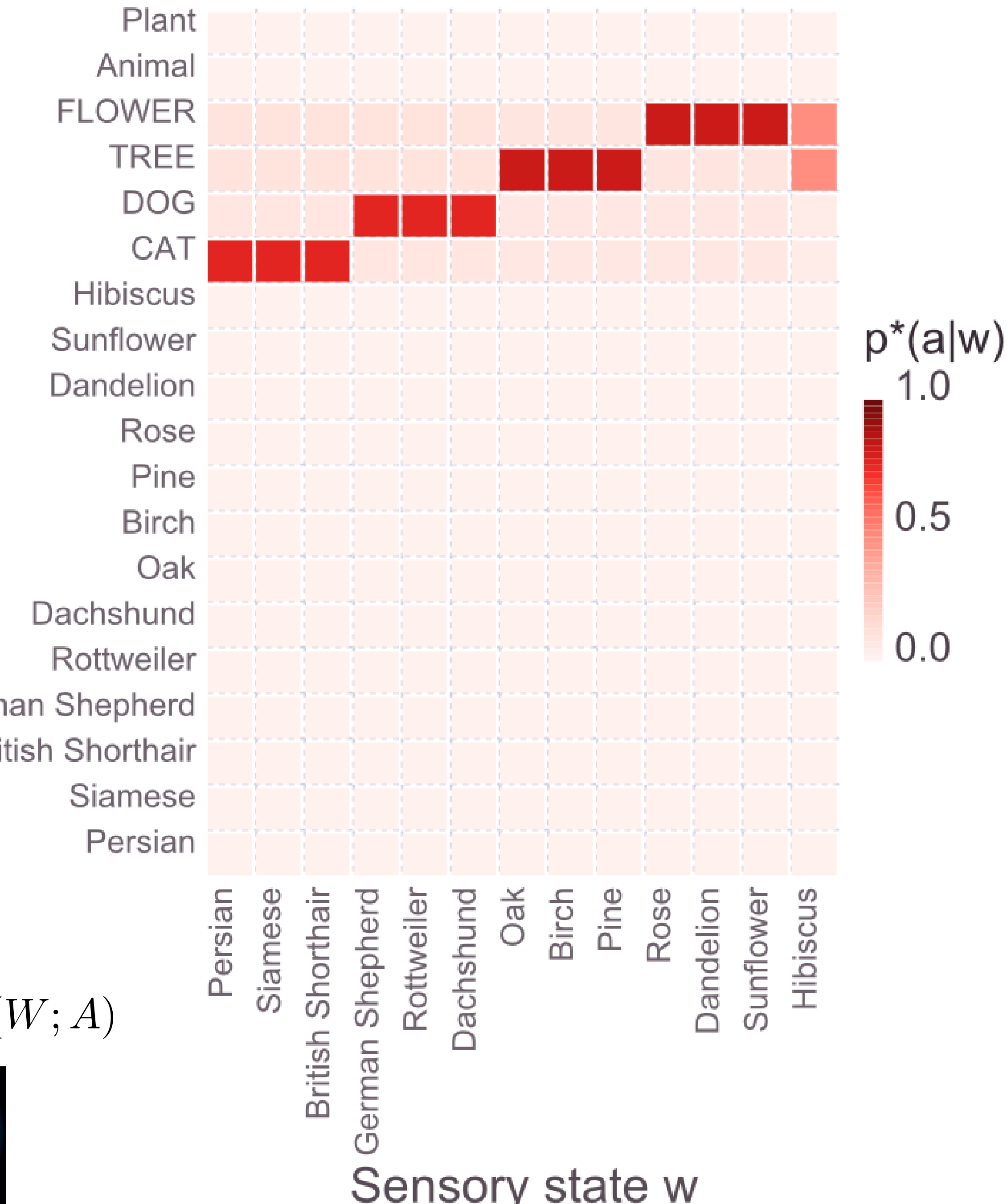
β	10	bits/€
I	3.7	bits
E[U]	3	€



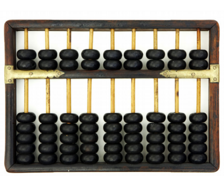


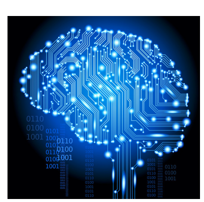


β	1.1	bits/€
I	0.9	bits
E[U]	1.8	€

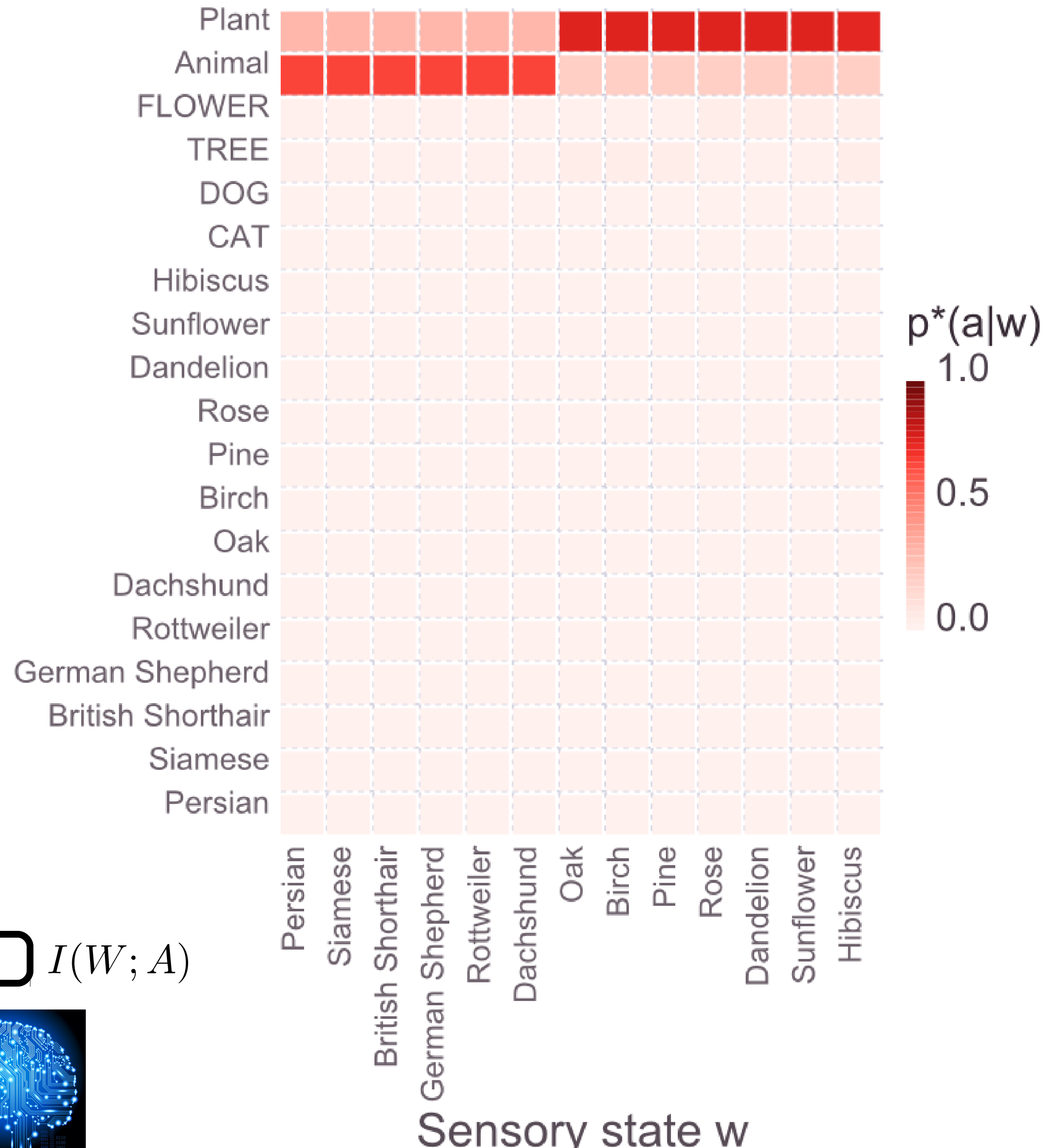


$I(W;A)$

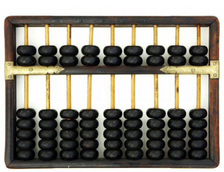


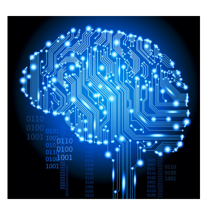


β	0.7	bits/€
I	0.2	bits
E[U]	1.2	€



$I(W;A)$

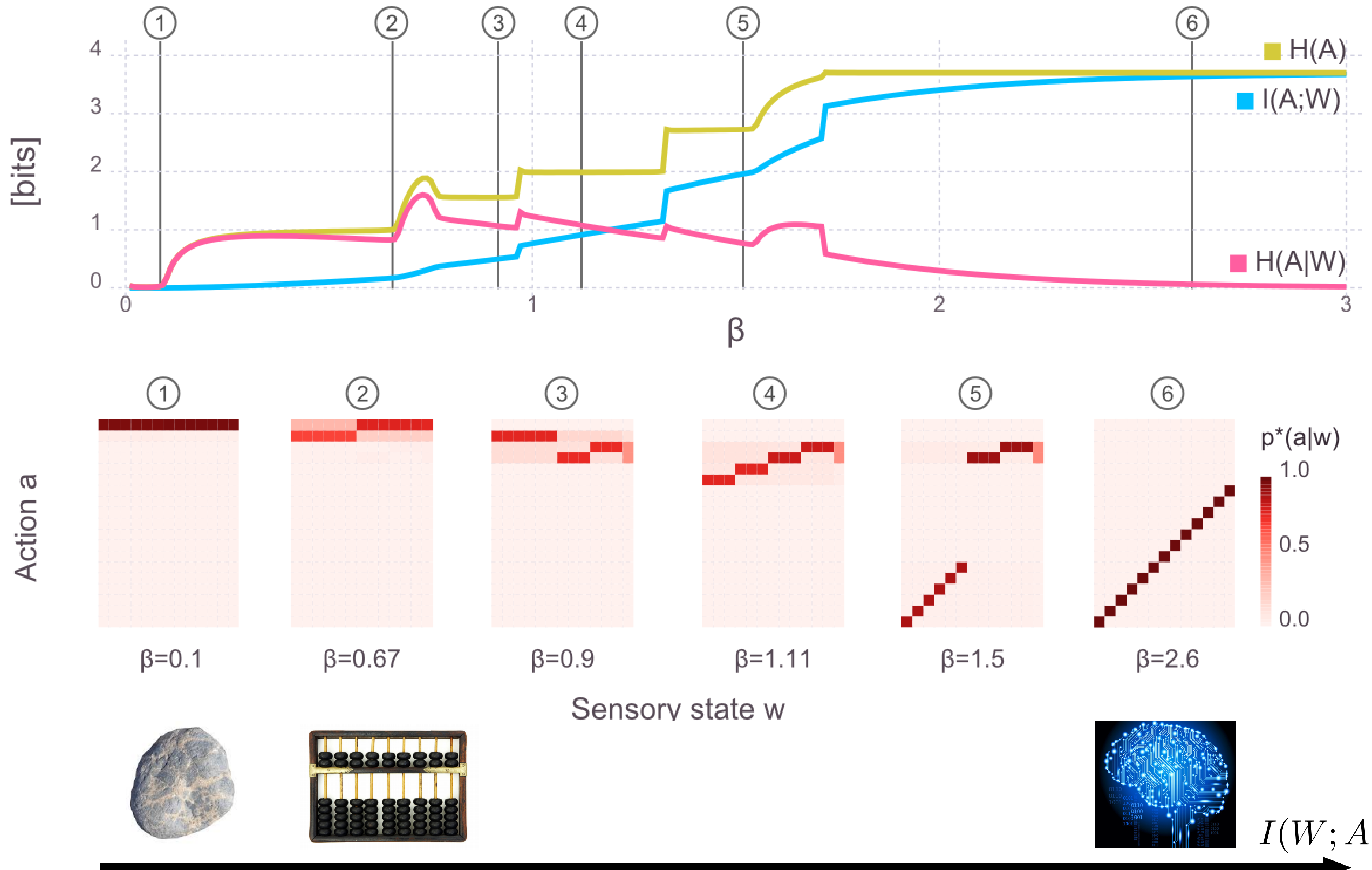




β	0.1	bits/€
I	0	bits
E[U]	0.9	€

$I(W;A)$

Emergence of natural levels of abstraction



Sensing for acting

Perception and action

World state w (not directly accessible), percept x, action a

$$w \longrightarrow x \longrightarrow a$$

Perception

Action

$$p(w|x) \longrightarrow U(w,a)$$

Perception and action

World state w (not directly accessible), percept x, action a

$$w \rightarrow x \rightarrow a$$

Action is decision-making

- Maximize utility under posterior belief over w

$$U(x,a) = \sum_{w} p(w|x)U(w,a)$$

Perception and action

World state w (not directly accessible), percept x, action a

$$w \rightarrow x \rightarrow a$$

Action is decision-making

- Maximize utility under posterior belief over w

$$U(x,a) = \sum_{w} p(w|x)U(w,a)$$

Classical: perception is inference

- Percept x should represent world-state w as faithfully as possible

$$p(w|x) = \frac{p(x|w)p(w)}{Z}$$

Perception and action

World state w (not directly accessible), percept x, action a

$$w \rightarrow x \rightarrow a$$

How to define the likelihood model $p(x|w)$?

Classical: problem independent of action-channel

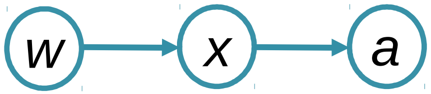
Classical: perception is inference

- Percept x should represent world-state w as faithfully as possible

$$p(w|x) = \frac{p(x|w)p(w)}{Z}$$

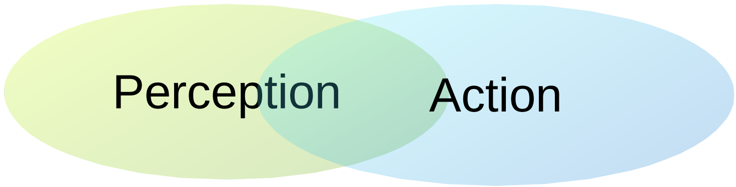
Perception and action

World state w (not directly accessible), percept x, action a



Information-theoretic bounded rationality:

- Trade off (gains in) expected utility against computational effort
- Two **coupled** stages of computation (“channels”):



Coupling perception and action

$$\underset{p(x|w),p(a|x)}{\arg\max}\ \mathbf{E}_{p(w,x,a)}[U(w,a)] - \frac{1}{\beta_1} I(W;X) - \frac{1}{\beta_2} I(X;A)$$

Set of self-consistent solutions:

$$p^*(x|w) = \frac{1}{Z(w)} p(x) \exp\left(\beta_1 \Delta F_{\text{ser}}(w,x)\right)$$

$$p(x) = \sum_w p(w) p^*(x|w)$$

$$p^*(a|x) = \frac{1}{Z(x)} p(a) \exp\left(\beta_2 \sum_w p(w|x) U(w,a)\right)$$

$$p(a) = \sum_{w,x} p(w) p^*(x|w) p^*(a|x),$$

Coupling perception and action

$$\underset{p(x|w),p(a|x)}{\arg\max}\ \mathbf{E}_{p(w,x,a)}[U(w,a)] - \frac{1}{\beta_1} I(W;X) - \frac{1}{\beta_2} I(X;A)$$

Set of self-consistent solutions:

$$\boxed{p^*(x|w) = \frac{1}{Z(w)} p(x) \exp\left(\beta_1 \Delta F_{\text{ser}}(w,x)\right)}$$

$$\Delta F_{\text{ser}}(w,x) := \mathbf{E}_{p^*(a|x)}[U(w,a)] - \frac{1}{\beta_2} D_{\text{KL}}(p^*(a|x)||p(a))$$

Well-defined likelihood model $p(x|w)$

- Maximizes downstream utility-computation trade-off (free energy)
- Tight coupling between perception and action!

Coupling perception and action

$$\underset{p(x|w),p(a|x)}{\arg\max}\ \mathbf{E}_{p(w,x,a)}[U(w,a)] - \frac{1}{\beta_1} I(W;X) - \frac{1}{\beta_2} I(X;A)$$

Set of self-consistent solutions:

$$p^*(x|w) = \frac{1}{Z(w)} p(x) \exp\left(\beta_1 \Delta F_{\text{ser}}(w,x)\right)$$

Bounded-optimal perception should extract the most **relevant information** (for efficient acting) rather than allowing to predict w as well as possible!

Coupling perception and action

$$\arg\max_{p(x|w),p(a|x)} \mathbf{E}_{p(w,x,a)}[U(w,a)] - \frac{1}{\beta_1} I(W;X) - \frac{1}{\beta_2} I(X;A)$$

Set of self-consistent solutions:

$$\boxed{p^*(a|x) = \frac{1}{Z(x)} p(a) \exp\left(\beta_2 \sum_w p(w|x) U(w,a)\right)}$$

The action-part of the system $p(a|x)$

- Maximizes posterior expected utility in a bounded rational fashion

$$U(x,a) = \sum_w p(w|x) U(w,a)$$

Predator-Prey example

Three groups of animals

- Small: prey, can’t hear well
- Medium: prey, can hear well
- Large: predators

Three basic actions

- Ambush: does not produce noise
- Sneak-up: does not work for medium-sized animals
- Flee: only sensible action when facing large animal

Predator-Prey example

Three groups of animals

- Small: prey, can’t hear well
- Medium: prey, can hear well
- Large: predators

Three basic actions

- Ambusl
- Sneak-
- Flee: or

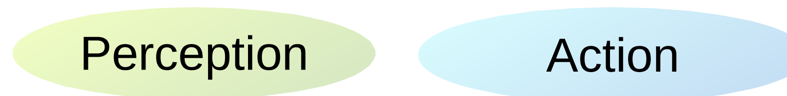
Utility depends on:
- Predator or Prey?
- Small or medium-sized animal?
  - Which kind of small-sized animal?

Predator-Prey example

Design agents with limited perceptual capacity

- bounded-optimal perception $p^*(x|w)$

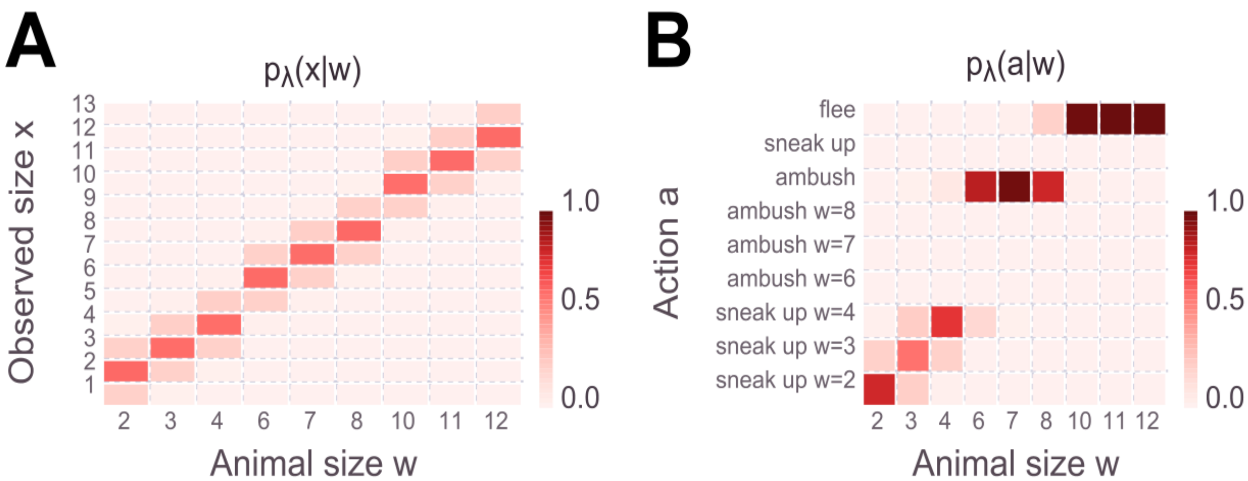
- hand-crafted likelihood model $p_\lambda(x|w)$



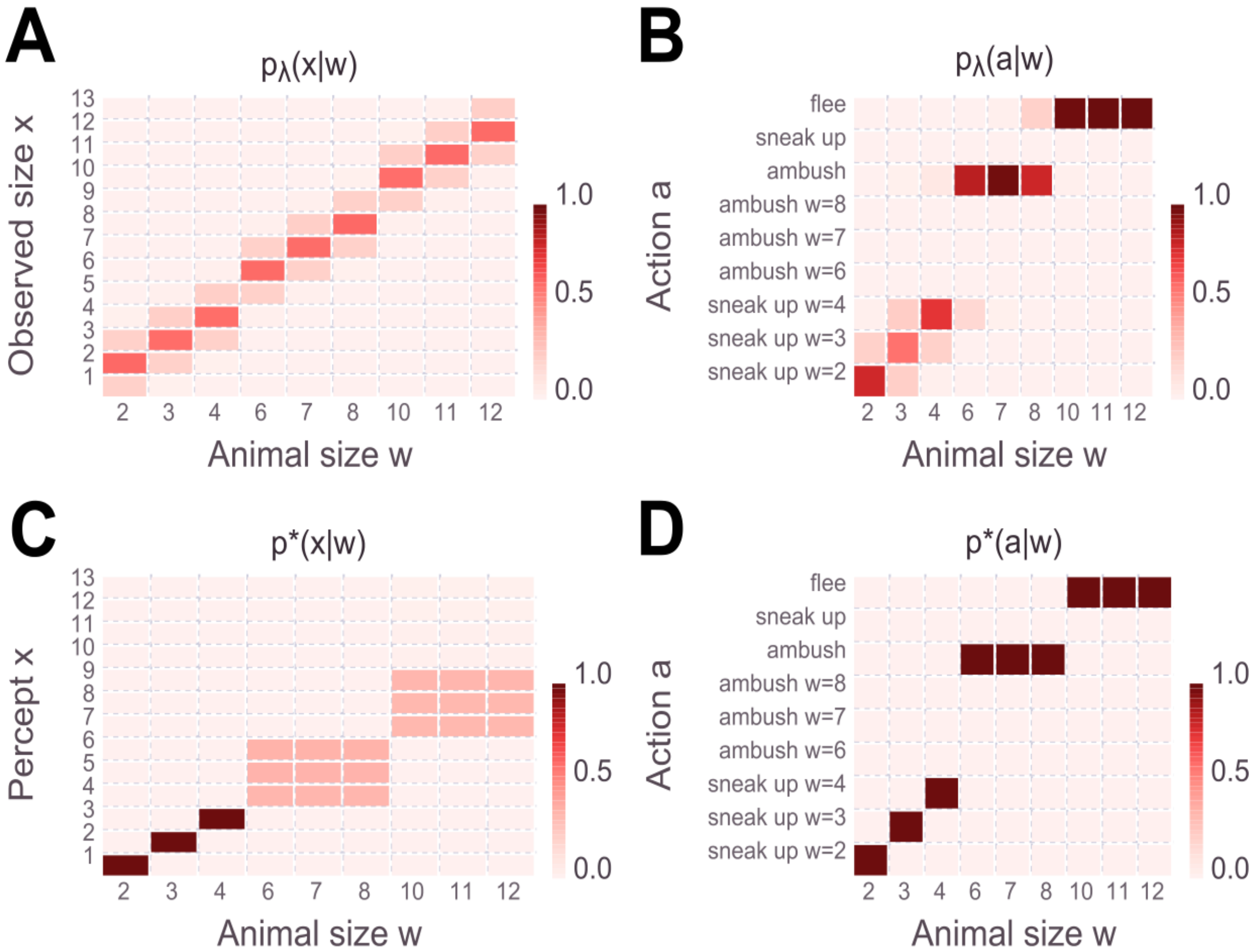
Hand-crafted model is designed to predict actual animal size w from observed animal size x well

- Gaussian perception
- Perceptual noise governs channel capacity
- Amount of noise can be adjusted through precision λ

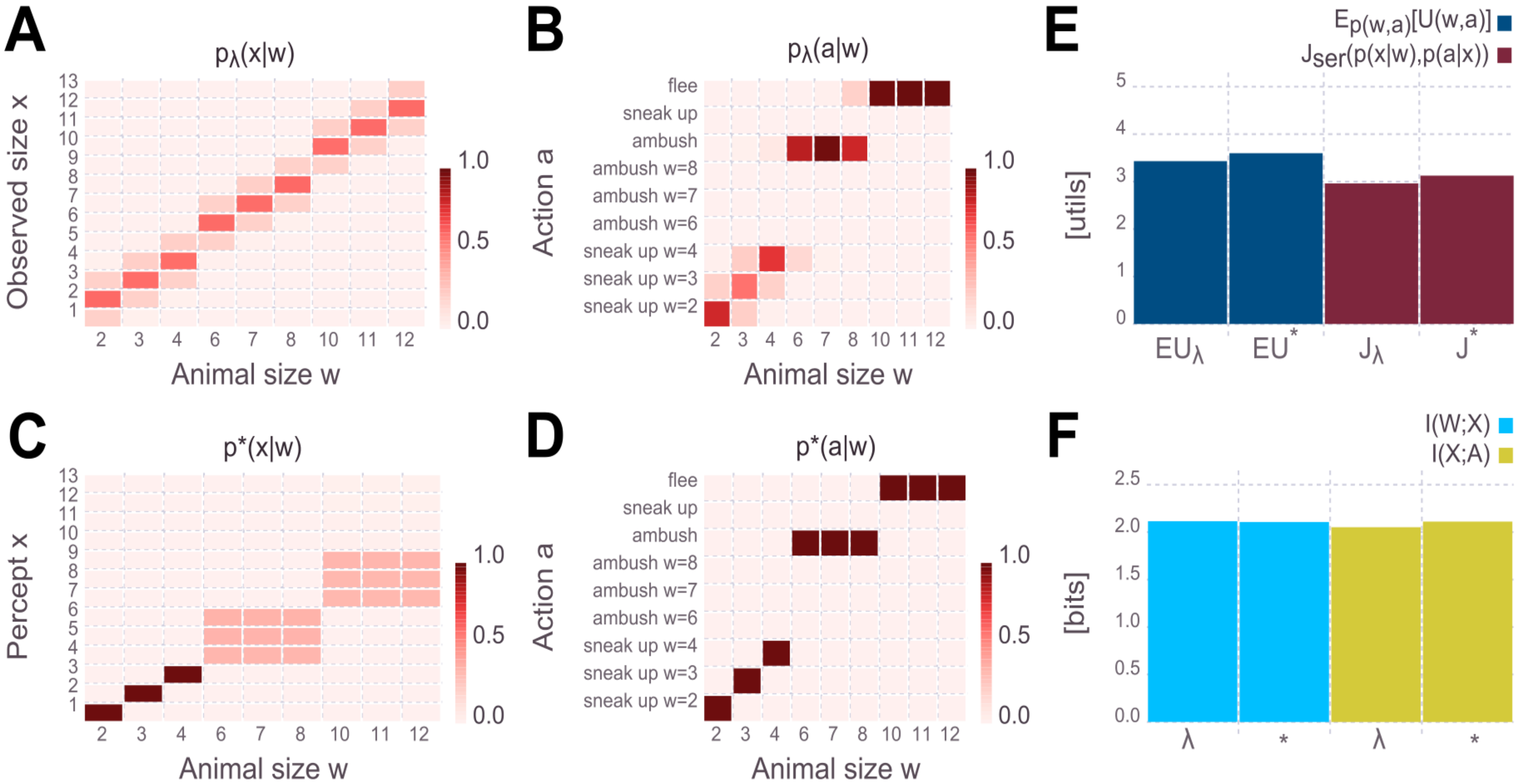
Large action capacity (good motor skills)



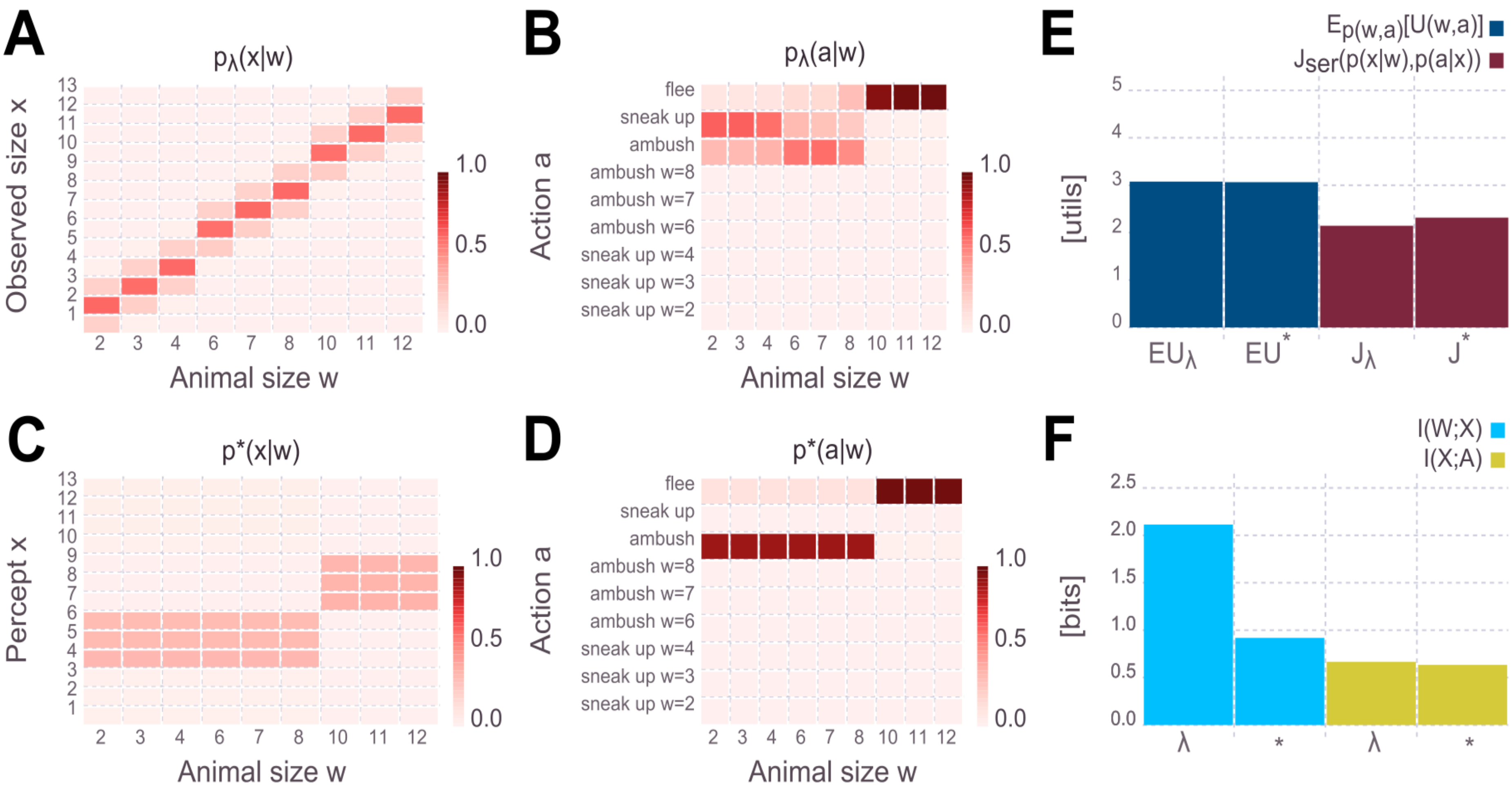
Large action capacity (good motor skills)



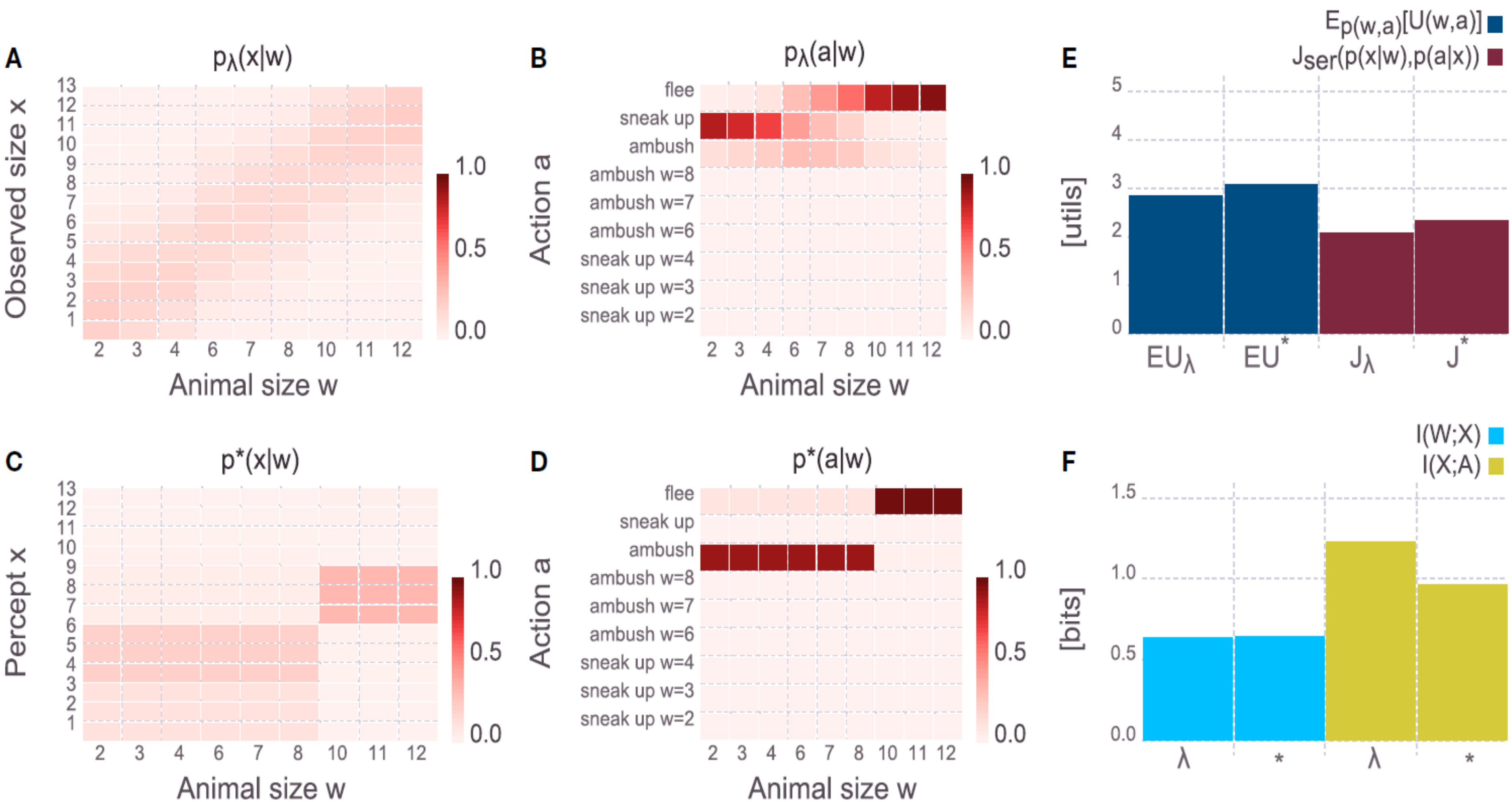
Large action capacity (good motor skills)



Low action capacity (bad motor skills)

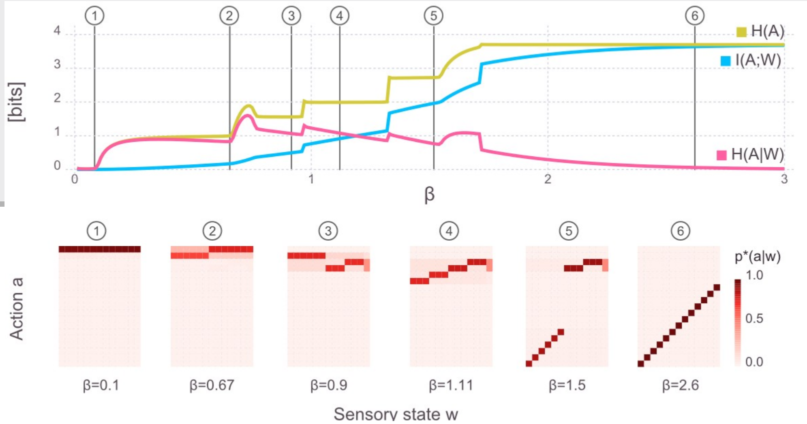


Low perceptual capacity (bad vision)



Hierarchies of abstraction

Parallel hierarchy

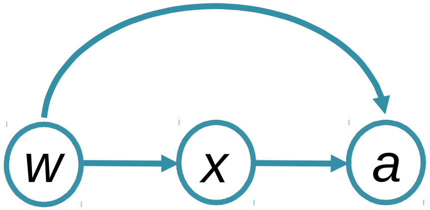


- Previously: Animal-Plant taxonomy
  - Inverse temperature β governs granularity of abstraction
  - Multiple levels of abstraction in parallel?

- Model-parameter hierarchy:
    - First step: “it’s an animal”
    - Second step: “it’s a cat”
  - First step reduces size of search space for second step
    - Distribution of information processing

Parallel hierarchy

Observation w, model x, action a

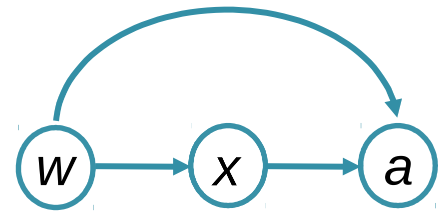


Two-level hierarchy

- High-level (priors over actions): $p(a|x)$
- Low-level (posteriors over actions): $p(a|w,x)$
- + model selector (weighting): $p(x|w)$

Model x narrows down search space for low-level decision

Parallel hierarchy



Observation w, model x, action a

Two-level hierarchy

- High-level (priors over actions): $p(a|x)$
- Low-level (posteriors over actions): $p(a|w,x)$
- + model selector (weighting): $p(x|w)$

All three parts of the hierarchy are well-defined through the solution to:

$$\underset{p(x|w),p(a|w,x)}{\arg\max}\ \mathbf{E}_{p(w,x,a)}[U(w,a)] - \frac{1}{\beta_1} I(W;X) - \frac{1}{\beta_3} I(W;A|X)$$

Parallel hierarchy

Optimality principle:

$$\underset{p(x|w),p(a|w,x)}{\arg\max}\ \mathbf{E}_{p(w,x,a)}[U(w,a)] - \frac{1}{\beta_1} I(W;X) - \frac{1}{\beta_3} I(W;A|X)$$

Self-consistent solutions:

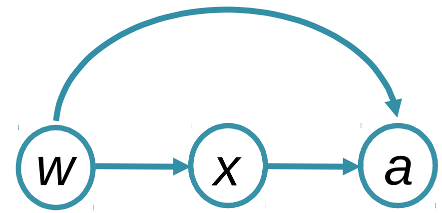
$$p^*(x|w) = \frac{1}{Z(w)} p(x) \exp\left(\beta_1 \Delta F_{\text{par}}(w,x)\right)$$

$$p(x) = \sum_w p(w) p^*(x|w)$$

$$p^*(a|w,x) = \frac{1}{Z(w,x)} p^*(a|x) \exp\left(\beta_3 U(w,a)\right)$$

$$p^*(a|x) = \sum_w p(w|x) p^*(a|w,x),$$

Parallel hierarchy



Observation w, model x, action a

Two-level hierarchy

- High-level (priors over actions): $p(a|x)$
- Low-level (posteriors over actions): $p(a|w, x)$
- + model selector (weighting): $p(x|w)$

Emergent hierarchy, shaped by:
- the utility function $U(w, a)$
- the environment $p(w)$
- and the computational limitations β_1, β_3

Conclusions

Conclusions

- Rate-Distortion principle for decision-making
  - Inverse temperature governs granularity of abstraction
  - Emergence of natural levels of abstraction

- Sensing for acting
  - Perception should extract most relevant information

- Hierarchies of abstraction
  - Distribution of information processing load
  - Extraction of transferable knowledge (?)

More information

Today's talk (my part):

- http://tim.inversetemperature.net/research/
- Explore all examples (Jupyter notebooks):
  https://github.com/tgenewein/BoundedRationalityAbstractionAndHierarchicalDecisionMaking
- Read the full paper:
  Genewein T., Leibfried F., Grau-Moya J., Braun D.A. (2015):
  Bounded rationality, abstraction and hierarchical decision-making: an information-theoretic optimality principle
  Frontiers in Robotics and AI, DOI:10.3389/frobt.2015.00027

Pedro A Ortega

Daniel A Braun

Jordi Grau-Moya

Felix Leibfried

Zhen Peng

Thanks!

- Analytical solution

$$p^*(a|w) = \frac{1}{Z(w)} p_0(a) e^{\beta U(w,a)}$$

  - Still intractable (partition sum):

$$\mathrm{Z(w)} = \sum_a p_0(a) e^{\beta U(w,a)}$$

  - Descriptive framework (external point-of-view)

Rejection sampling scheme

- Constructive framework (internal point-of-view)

1. Draw sample $\tilde{a} \sim p_{0(a)}, \quad \tilde{u} \sim U(0,1)$
2. Accept if: $\tilde{u} \leq \frac{e^{\beta U(w,a)}}{e^{\beta U_{\max}(w)}}$
3. Otherwise reject $\tilde{a}$ and go back to 1.

- Guaranteed to produce samples from $p^*(a|w)$
- Expected number of rejections: $\#\text{samples} = e^{\beta U_{\max}(w)} / Z(w)$
  - β controls how many rejections "are allowed" (on average)

Future directions:

Continuous problems / parametric distributions, large-scale problems

- Sampling schemes

  Ortega, P.A., Braun, D.A., Dyer, J.S., Kim, K.-E., and Tishby, N.
  Information-Theoretic Bounded Rationality
  ArXiv:1512.06789, 2015

- Regularizer for (deep) neural networks

  Leibfried, F., Braun, D.A
  Bounded Rational Decision-Making in Feedforward Neural Networks
  UAI 2016, ArXiv:1602.08332

Sequential decision-making problems (reinforcement learning)

- Tishby, Polani, and others.

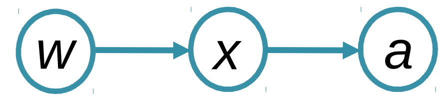
  Fox, R., Pakman, A., Tishby, N.
  G-Learning: Taming the Noise in Reinforcement Learning via Soft Updates
  ArXiv:1512.08562, 2015

- Modeling computational limitations and model uncertainty

  Grau-Moya, J., Leibfried, F., Genewein, T., Braun, D.A
  Planning with Information-Processing Constraints and Model Uncertainty in Markov Decision Processes
  ArXiv:1604.02080, 2016

Distribution of information processing

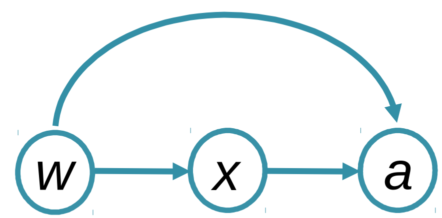
- Serial perception-action architecture:



  - Same information must pass both stages, must be paid-for twice

$$I(W;A) \leq \min\{I(W;X), I(X;A)\}$$

Hierarchical information processing:

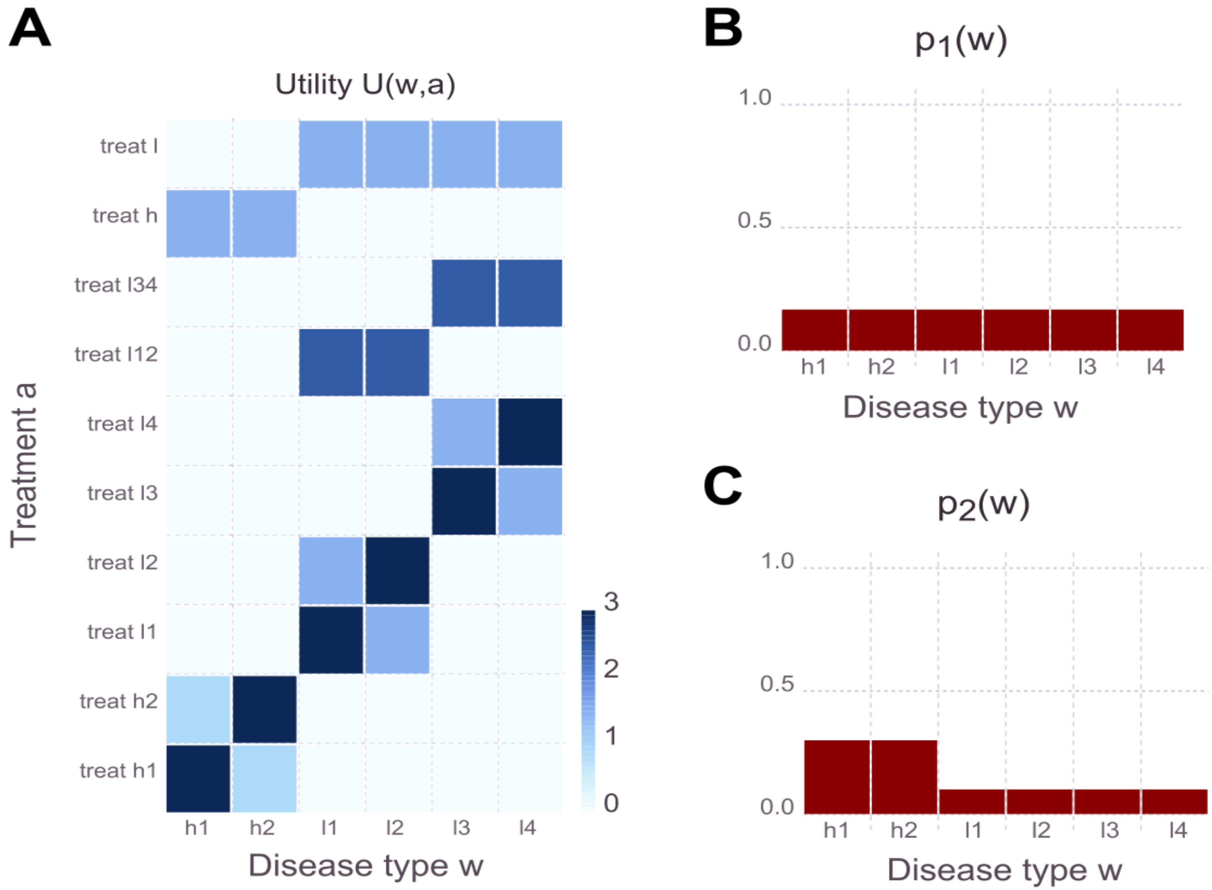


- Two-step reduction in uncertainty about a
- High-level stage: $\mathrm{p}(a|x)$
- Low-level stage: $\mathrm{p}(a|w,x)$

$$\underbrace{I(W,X;A)}_{\text{total reduction}} = \underbrace{I(X;A)}_{\text{high level}} + \underbrace{I(W;A|X)}_{\text{low level}}$$

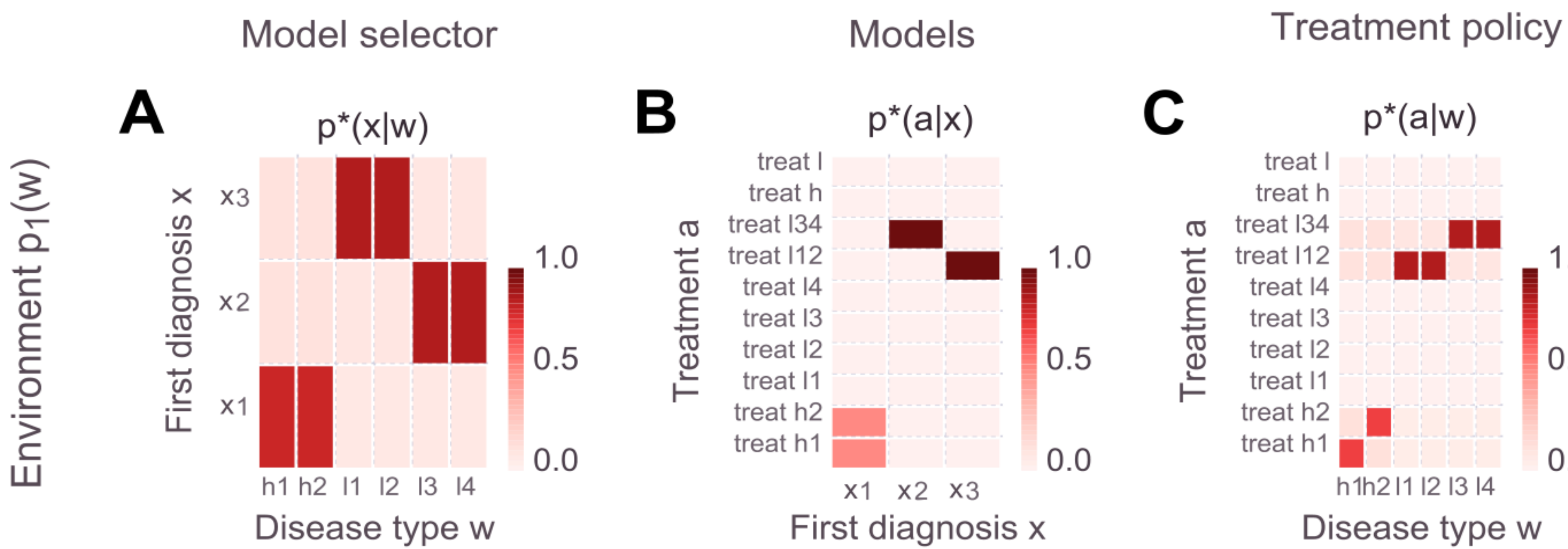
Medical system example

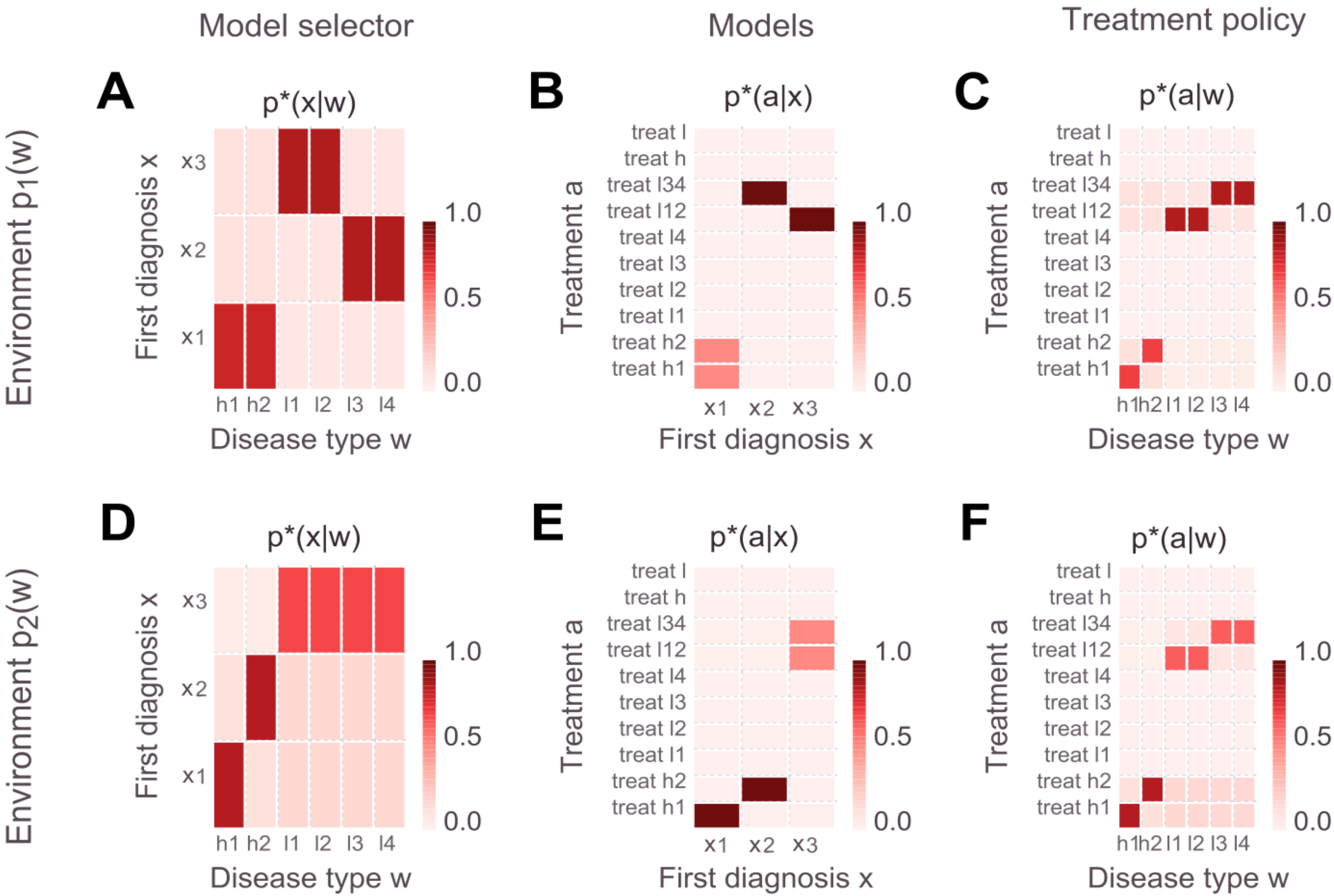
- Six different diseases
- Specific treatments and less effective general treatments



Diagnosis hierarchy

- Design of diagnosis hierarchy
  - Measurements by general practitioner
    - Standard measurements (cheap)
    - Leads to first diagnosis and sends patient to specialist
  - Specialists
    - Can perform additional measurements (expensive)
    - Maximally three kinds of specialists
    - Total budget for the specialists: 1 bit
      - E.g. only one specialist can have an MRI machine
      - Insufficient to diagnose every disease precisely
      - Rather: take the most important measurements
      - How should the system be constructed?


Model selector
Models
Treatment policy
A
B
C
Environment $p_1(w)$
p*(x|w)
First diagnosis x
x3
x2
x1
h1 h2 l1 l2 l3 l4
Disease type w
p*(a|x)
Treatment a
treat l
treat h
treat l34
treat l12
treat l4
treat l3
treat l2
treat l1
treat h2
treat h1
x1 x2 x3
First diagnosis x
p*(a|w)
1.0
0.5
0.0



Model selector
Models
Treatment policy
Environment p1(w)
Environment p2(w)
A
p*(x|w)
First diagnosis x
x3
x2
x1
h1 h2 l1 l2 l3 l4
Disease type w
1.0
0.5
0.0
B
p*(a|x)
Treatment a
treat l
treat h
treat l34
treat l12
treat l4
treat l3
treat l2
treat l1
treat h2
treat h1
x1 x2 x3
First diagnosis x
C
p*(a|w)
D
E
F