INFORMATION-THEORETIC BOUNDED RATIONALITY IN PERCEPTION-ACTION SYSTEMS

Tim Genewein

Sensorimotor Learning and Decision Making Group
Max Planck Institute for Intelligent Systems
Max Planck Institute for Biological Cybernetics

inversetemperature.net

tim.genewein@gmail.com

SENSORIMOTOR LEARNING AND DECISION-MAKING GROUP

Daniel A. Braun

MAX PLANCK INSTITUTE FOR INTELLIGENT SYSTEMS

Sensorimotor Learning

- Bayesian models
- Structure Learning
- Hierarchies of abstraction

Theory of Decision-Making

- Neuroeconomic principles
- Bounded rationality

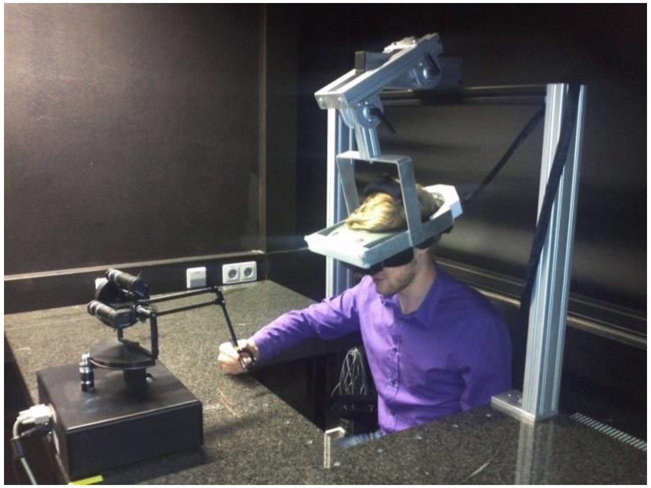
SENSORIMOTOR LEARNING AND DECISION-MAKING GROUP

Daniel A. Braun

Sensorimotor Learning
- Bayesian models
- Structure Learning
- Hierarchies of abstraction

Theory of Decision-Making
- Neuroeconomic principles
- Bounded rationality

Experiments in virtual-reality (psychophysics)



PhD: Structure Learning with Hierarchical Models for Computational Motor Control

SENSORIMOTOR LEARNING AND DECISION-MAKING GROUP

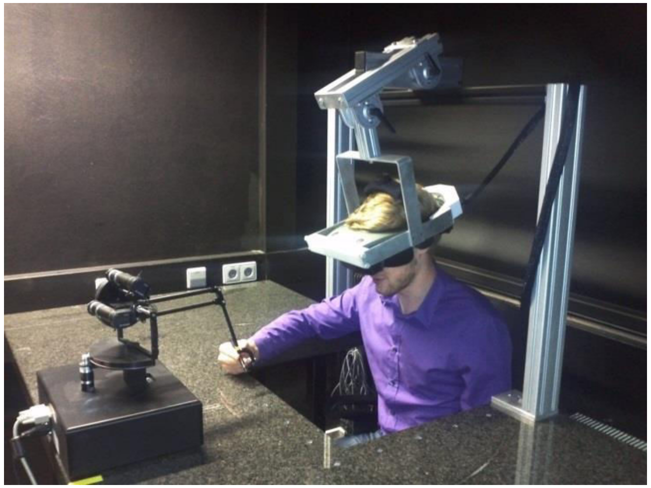
Daniel A. Braun

MAX PLANCK INSTITUTE FOR INTELLIGENT SYSTEMS

Sensorimotor Learning
- Bayesian models
- Structure Learning
- Hierarchies of abstraction

Theory of Decision-Making
- Neuroeconomic principles
- Bounded rationality

Experiments in virtual-reality (psychophysics)



PhD: Structure Learning with Hierarchical Models for Computational Motor Control

From July 2016:

Cognitive Systems Group

Corporate Research - Robert Bosch GmbH

BOSCH

Outline

- Information-theoretic bounded rationality
  - Free energy minimization
  - Lossy compression and emergence of levels of abstraction

- Perception for Action
  - Coupling perception and action
  - Likelihood function synthetization
  - Illustrative example

Utility maximization

Goal:

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

Easy…

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

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 $U(w, a)$

Take the process of computation into account

- Modified optimality principle
- 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 w
- Any change of behavior requires **computation**
- Limit the "amount of change" in behavior

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

Information-theoretic bounded rationality

Trade-off:

- Large expected utility
- Low computational effort

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

- Mathematically equivalent to minimization of **free energy** difference
- Also, deep conceptual ties – “the physics of computation”
  - Ortega, Braun 2013, *Thermodynamics as a theory of decision-making with information-processing cos*ts, Royal Society 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 … 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

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

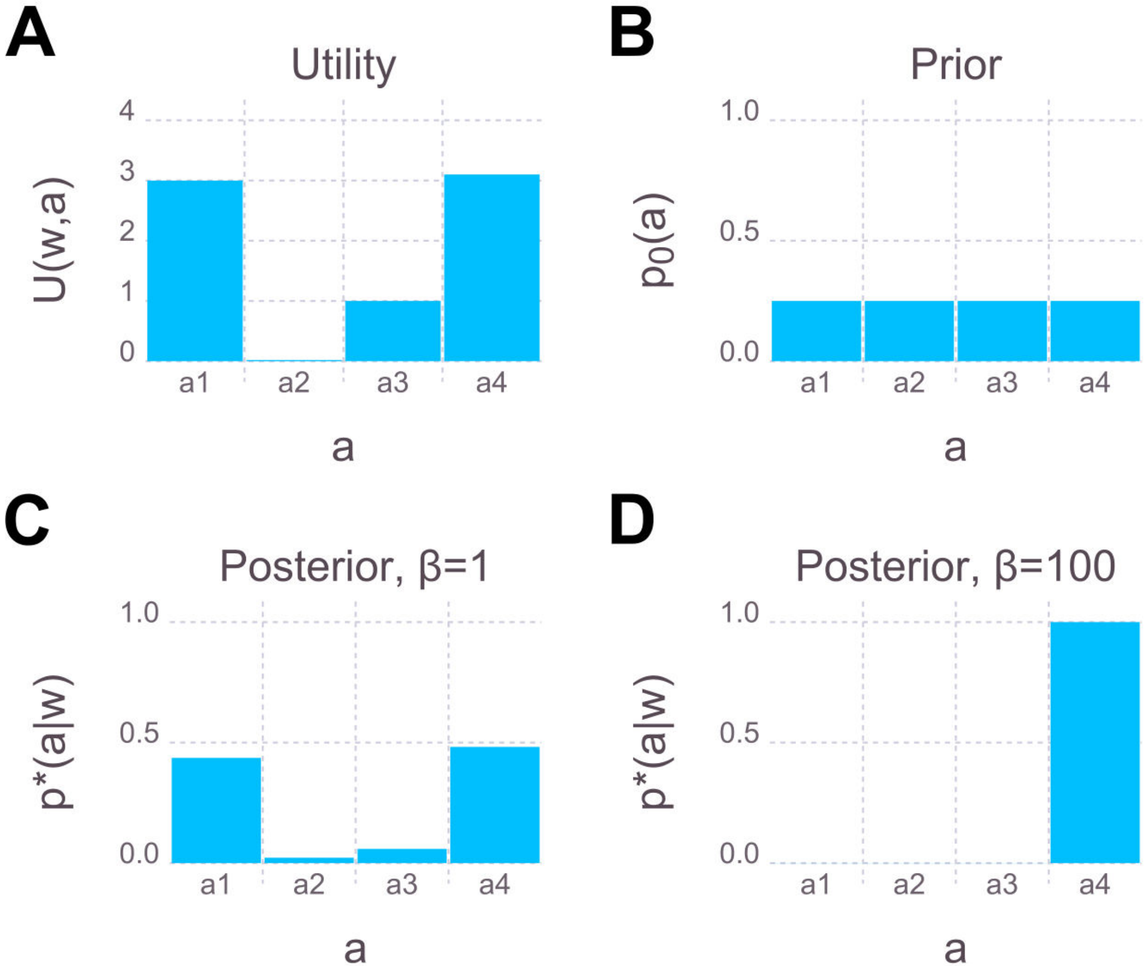
$$p^*(a|w) = \frac{1}{Z} p_0(a) e^{\beta U(w,a)}$$

<u>Special case – Bayes' rule:</u>

$$U(w,a) = \log q(w|a), \qquad \beta = 1$$

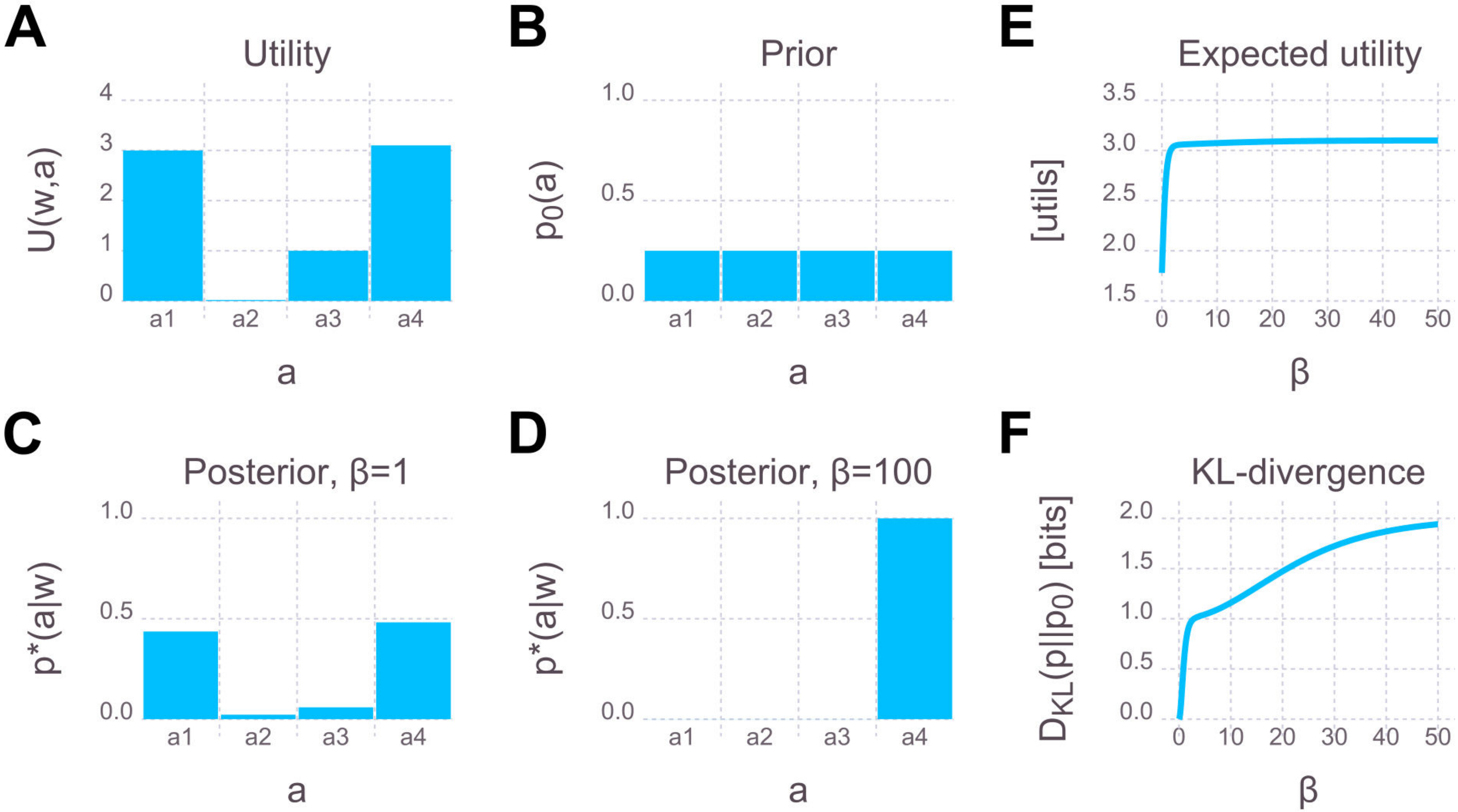
$$p^*(a|w) = \frac{q(w|a)p_0(a)}{Z}$$

Example: grasping movement



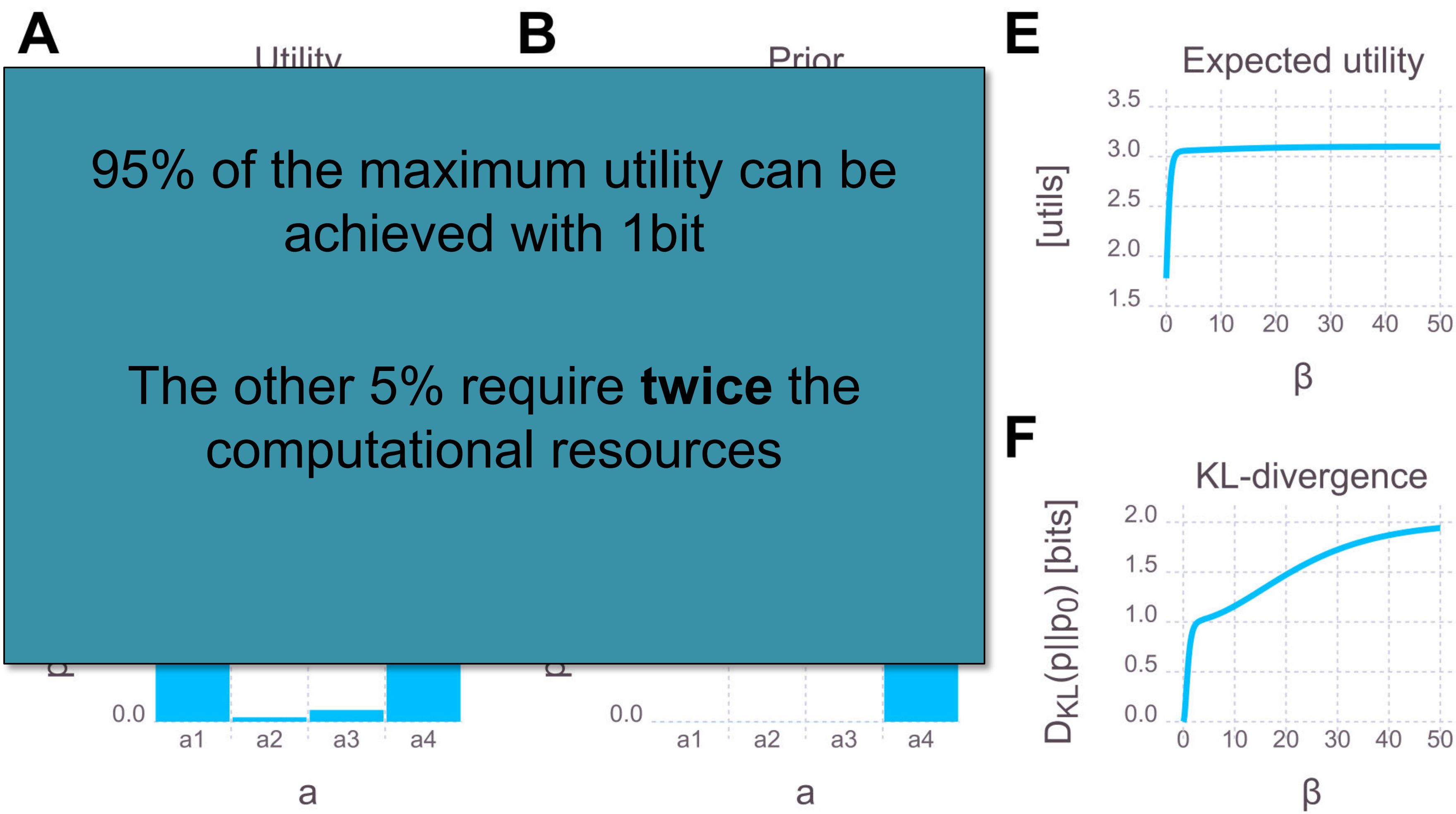
Example: grasping movement

- β governs computational resources



Example: grasping movement

- β governs computational resources



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

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

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)

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: $\#\mathrm{samples} = e^{\beta U_{\max}(w)} / Z(w)$
  - β controls how many rejections “are allowed” (on average)

Outline

- Information-theoretic bounded rationality
  - Free energy minimization
  - Lossy compression and emergence of levels of abstraction

- Perception for Action
  - Coupling perception and action
  - Likelihood function synthetization
  - Illustrative example

Multiple world-states

Trade off large utility against low computational effort

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

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

- Mathematically equivalent to **rate-distortion** problem
  - → Lossy compression
  - Channel from observations 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 $p_0(a)$? → the marginal
- 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:

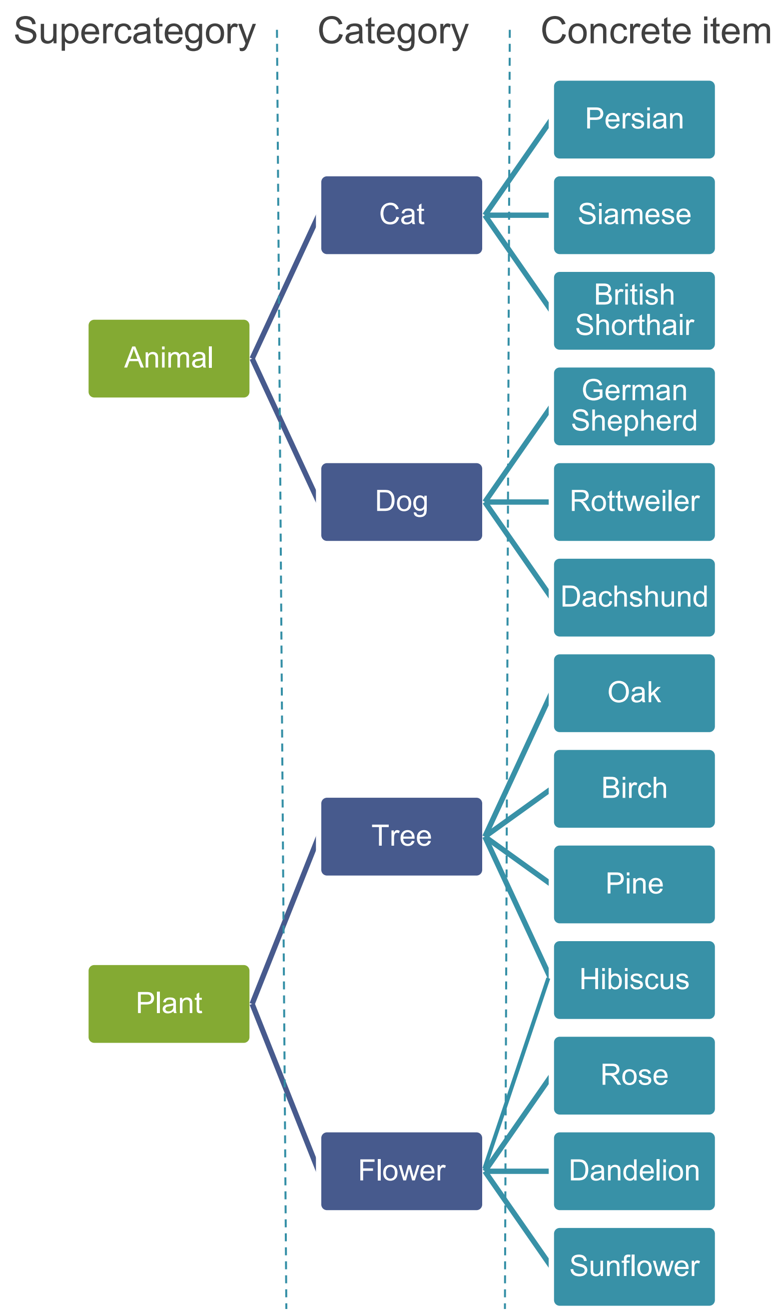
$$\boxed{\begin{aligned} p^*(a|w) &= \frac{1}{Z} p(a) e^{\beta U(w,a)} \\ p(a) &= \sum_w p(w) p^*(a|w) \end{aligned}}$$

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

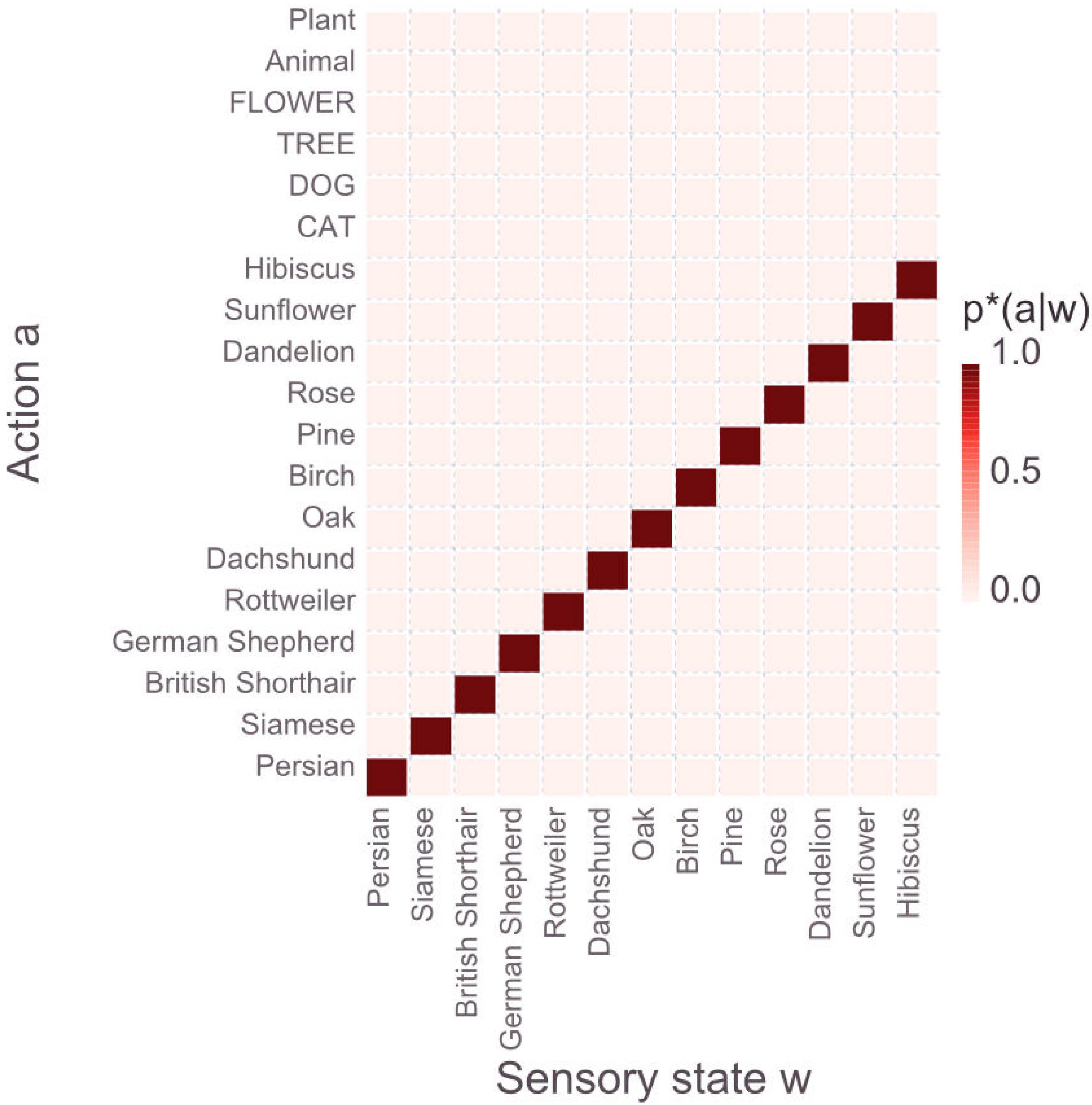
Toy Example

- Sensory state
  $w \in \{\text{concrete items}\}$
- Action $a \in \{\text{concrete items},$
  $\text{categories}, \text{supercategories}\}$

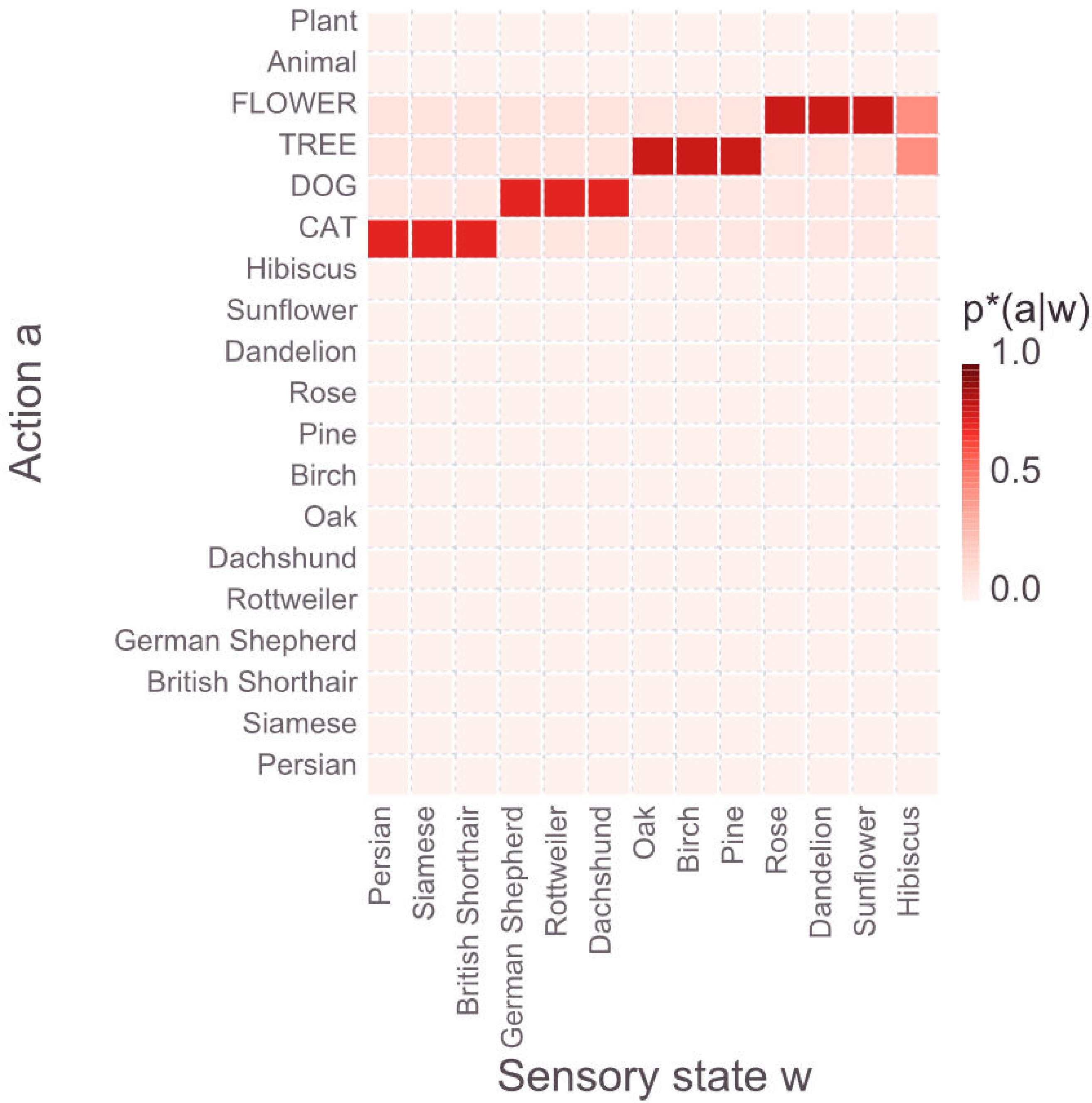
- Rewards/Utilities:
  - 3€ if concrete item correct
  - 2.2€ if category correct
  - 1.6€ if supercategory correct



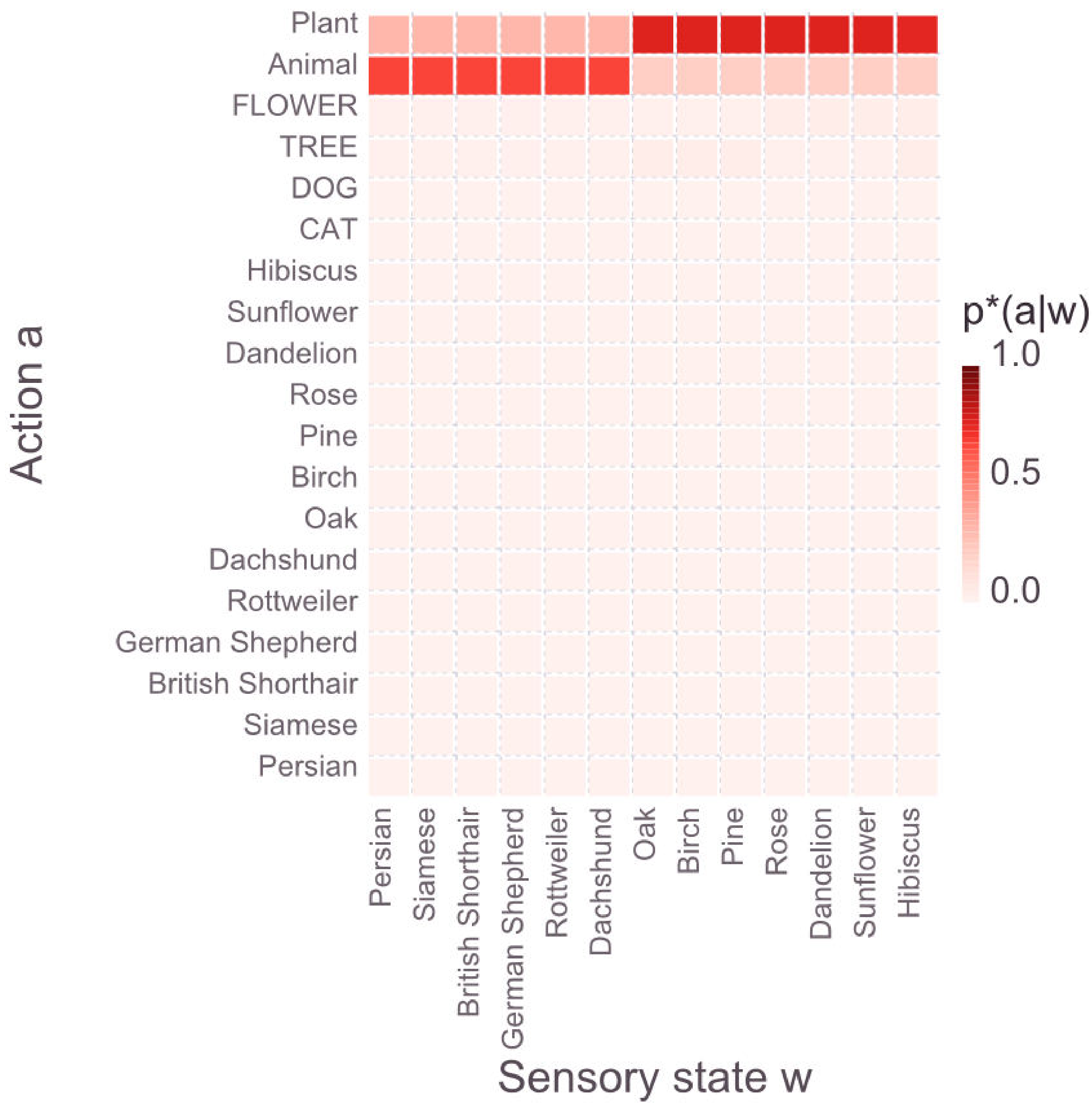
β	10	$[\text{bits}/€]$
I	3.7	[bits]
$\mathbf{E}[U]$	3	[€]



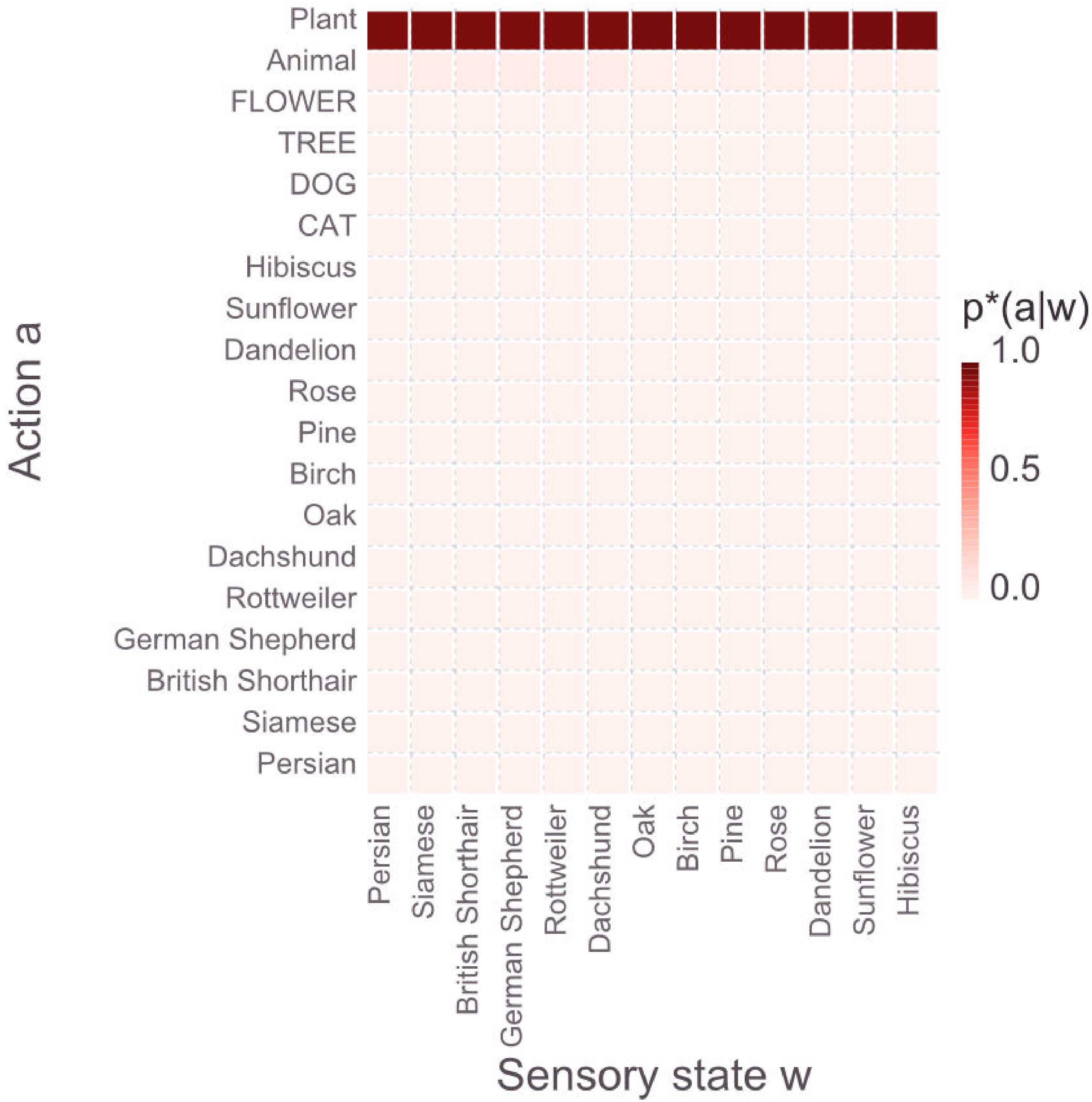
β	1.11	$\left[{}^{\text{bits}}/_{€}\right]$
I	0.9	[bits]
$\mathbf{E}[U]$	1.8	[€]



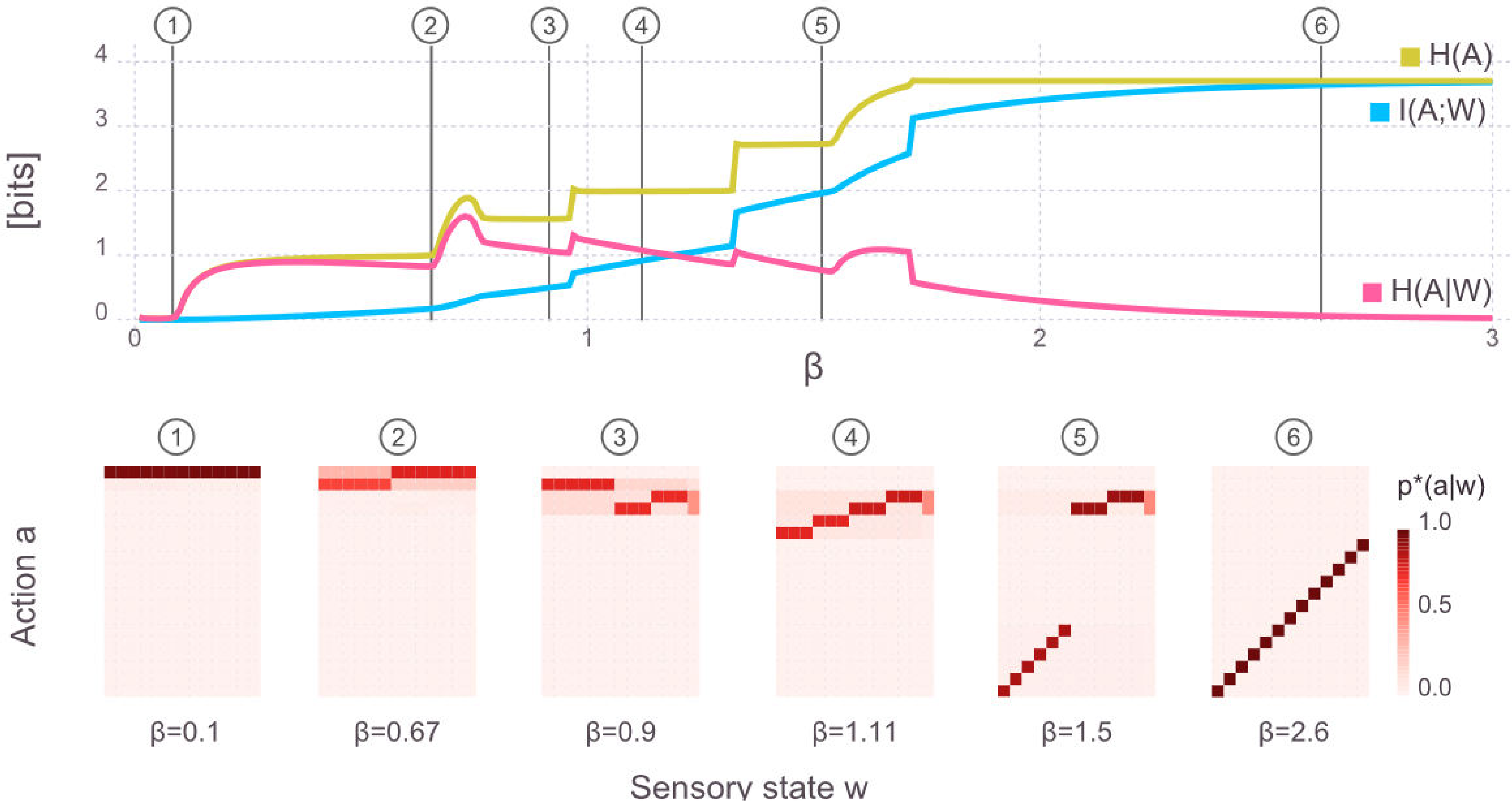
β	0.67	$\left[{}^{\text{bits}}/_{€}\right]$
I	0.18	[bits]
$\mathbf{E}[U]$	1.2	[€]



β	0.1	$[\text{bits}/\text{€}]$
I	0	[bits]
$\mathbf{E}[U]$	0.86	[€]



Emergence of natural levels of abstraction



Summary – bounded rationality

Trade off large utility against low computational cost

Abstractions are induced through limitations in information processing capabilities

- Levels of abstraction are formed through the structure of the utility-function

Extensions:

- Modelling perception-action systems with the principle
- Modelling several layers of abstraction in parallel (not in this talk)

Outline

- Information-theoretic bounded rationality
  - Free energy minimization
  - Lossy compression and emergence of levels of abstraction
- Perception for Action
  - Coupling perception and action
  - Likelihood function synthetization
  - Illustrative example

Perception and action

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

$$w \rightarrow x \rightarrow a$$

Classical: perception is inference

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

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

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

$$w \rightarrow x \rightarrow a$$

Classical: perception is inference

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

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

How to define the likelihood model $p(x|w)$?
Classical: problem independent of action-channel

Perception for action

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

$$w \to x \to a$$

Information-theoretic bounded rationality:

- Trade off (gains in) expected utility against computational effort

Two coupled stages of computation (“channels”):

- Perception: stochastic mapping from w to x
- Action: stochastic mapping from x to a

Trading off utility against information processing cost

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

Set of self-consistent solutions:

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

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

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

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

Trading off utility against information processing cost

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

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!

Trading off utility against information processing cost

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

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!

Trading off utility against information processing cost

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

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

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

Outline

- Information-theoretic bounded rationality
  - Free energy minimization
  - Lossy compression and emergence of levels of abstraction
- Perception for Action
  - Coupling perception and action
  - Likelihood function synthetization
  - Illustrative example

Predator-Prey example

Three groups of animals

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

Three basic actions

- Ambush: works equally well on small and medium-sized animals
- Sneak-up: works well on small animals
- Flee: only sensible actions for large animals

Predator-Prey example

Three groups of animals

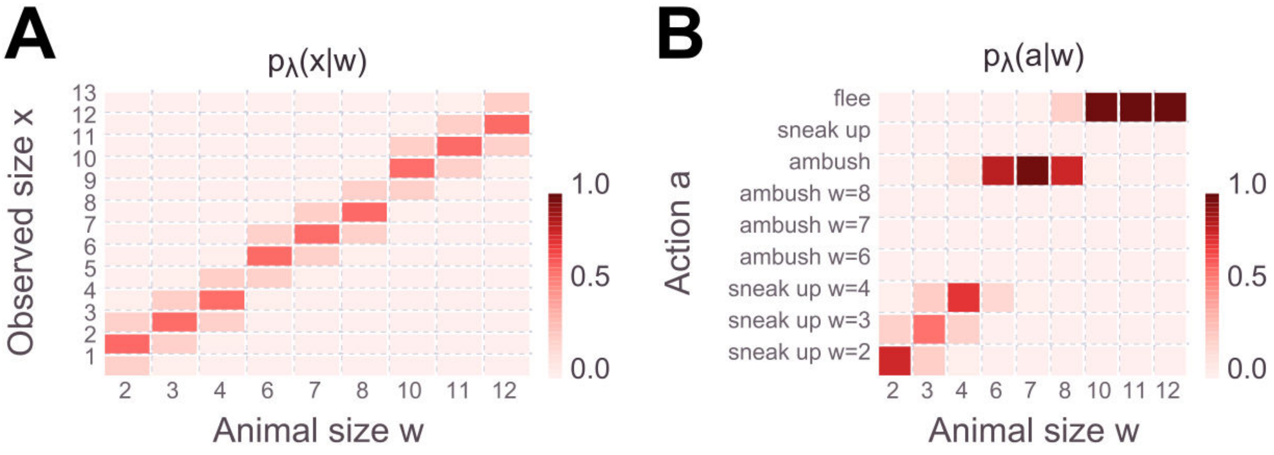
- Small: prey, can’t hear well
- Medium: prey, can hear well
- Large: predators

Three basic actions

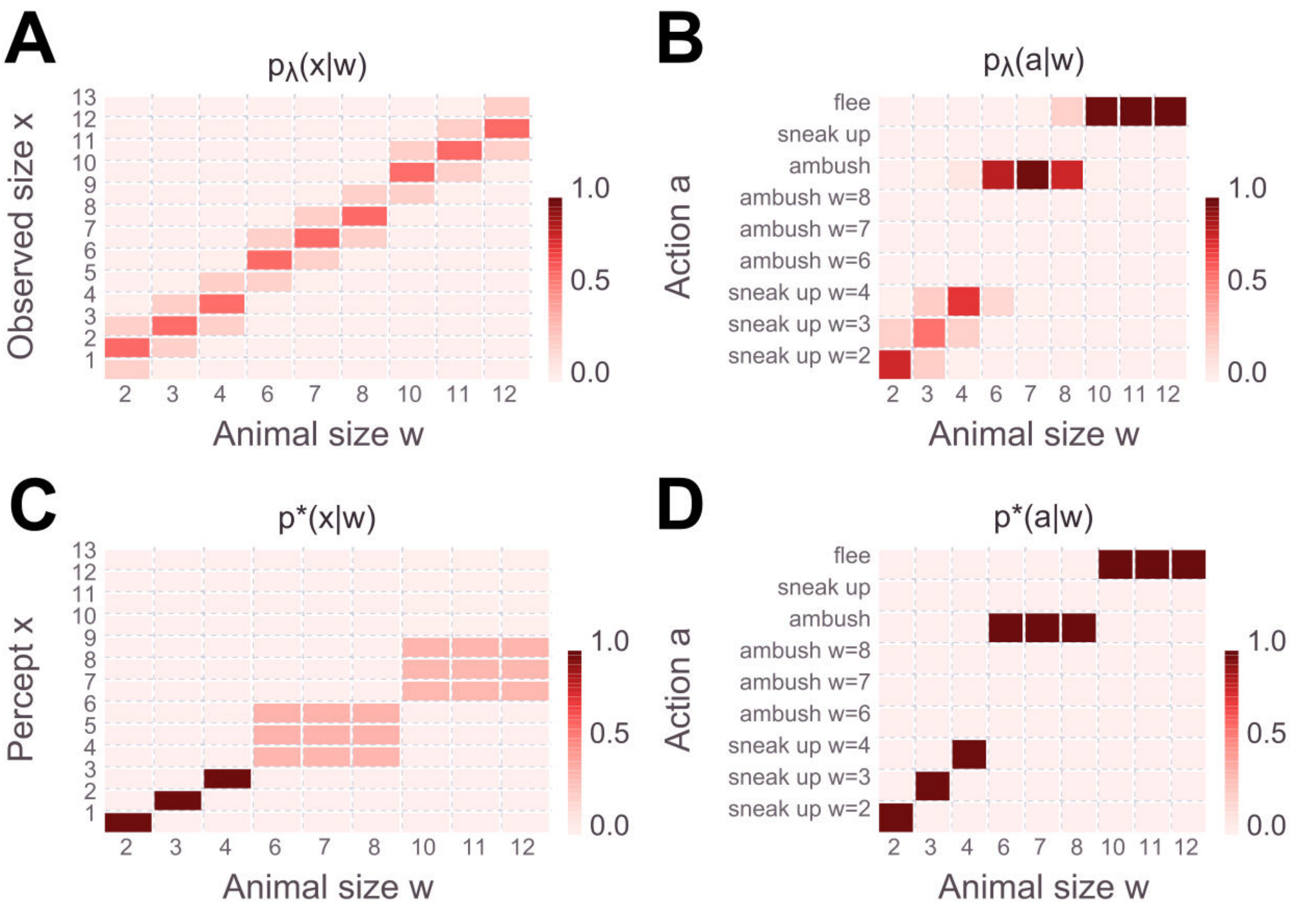
- Ambush: works equally well on small and medium-sized animals
- Sneak-up: works well on small animals
- Flee: only sensible actions for large animals

- Design agent with limited perceptual capacity
  - Compare bounded-optimal perception against hand-crafted likelihood model
  - Hand-crafted model is designed to predict actual animal size w from observed animal size x well

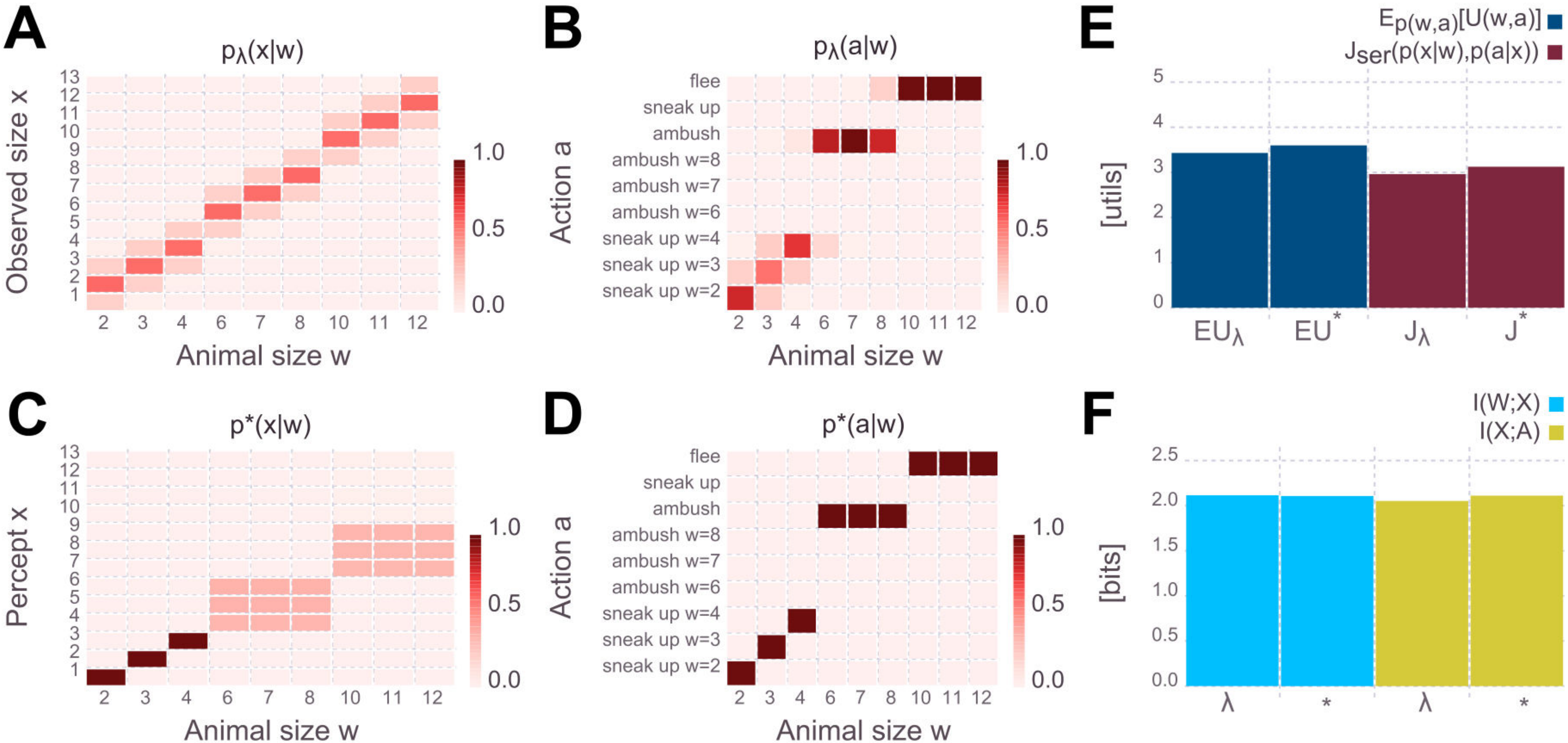
Large action capacity (good motor skills)



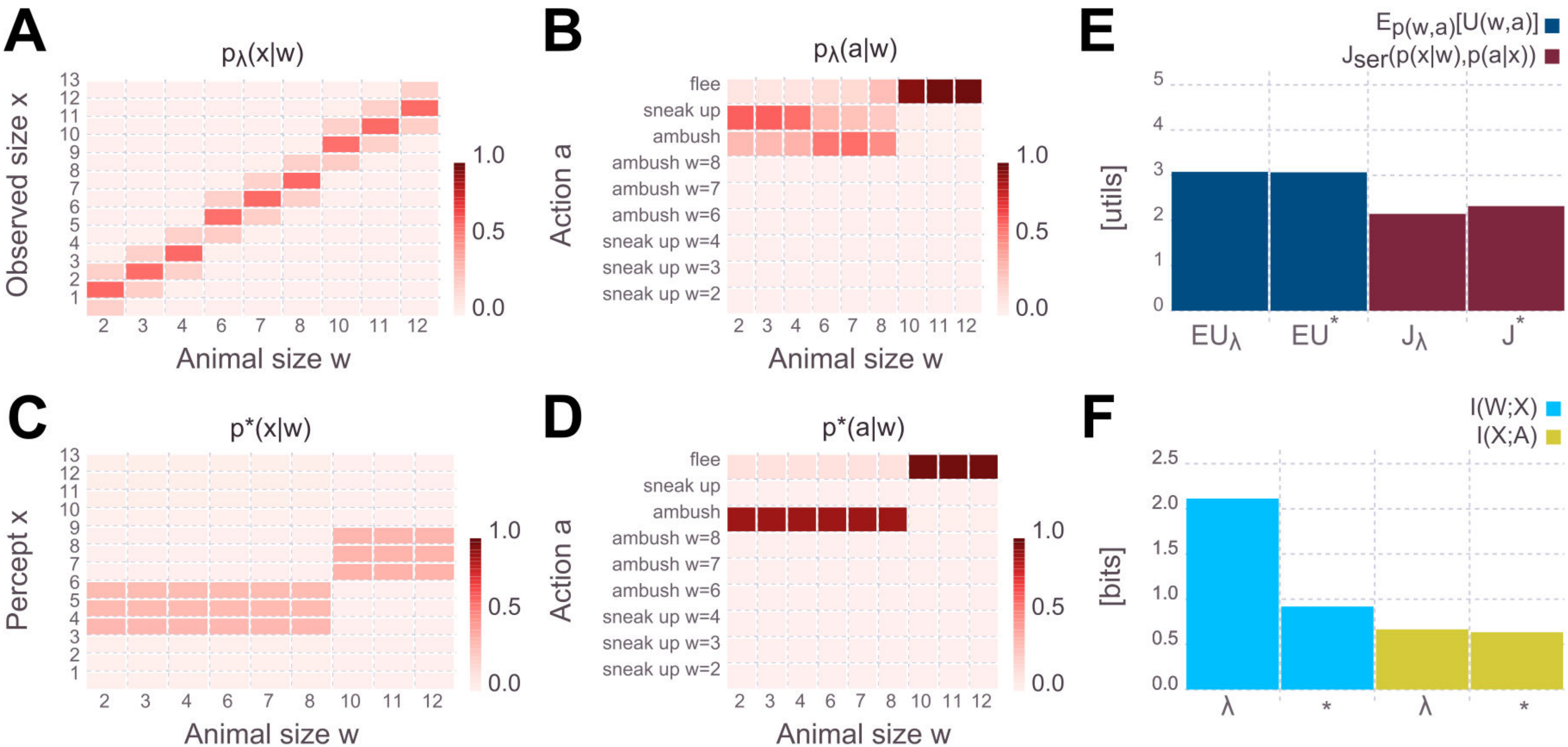
Large action capacity (good motor skills)



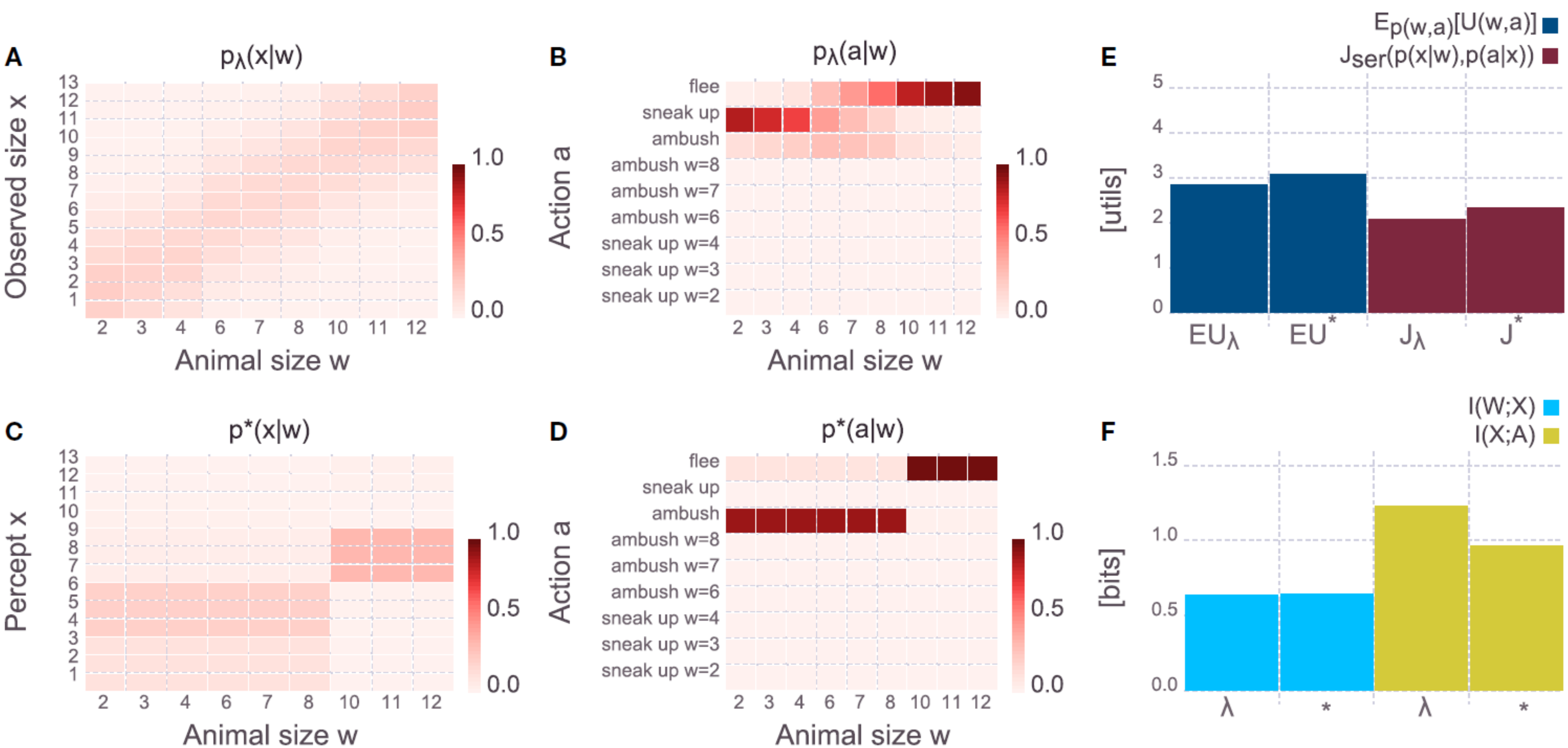
Large action capacity (good motor skills)



Low action capacity (bad motor skills)



Low perceptual capacity (bad vision)



Conclusions

Information-theoretic optimality principle

- Limited computational resources are part of the optimization problem
  - Formalization of limited resources as KL-divergence
  - Trade off gains in utility against cost of computation

- Mathematical and conceptual relations to:
  - Thermodynamics, Statistical mechanics
  - Information theory, lossy compression
  - Path integral control, variational Bayes, relative entropy policy search, KL-control, G-learning, …

- Modelling perception-action system
  - Bounded-optimal perception is tightly coupled to action
  - Likelihood function synthetization

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

More information

Today's talk:

- http://tim.inversetemperature.net/research/
- Explore all examples (Jupyter notebooks):
  https://github.com/tgenewein/BoundedRationalityAbstractionAndHierarchicalDecisionMaking
- 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

Information-theoretic bounded rationality

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

- Pedro Ortega's website: http://www.adaptiveagents.org/freeenergy

Thanks!