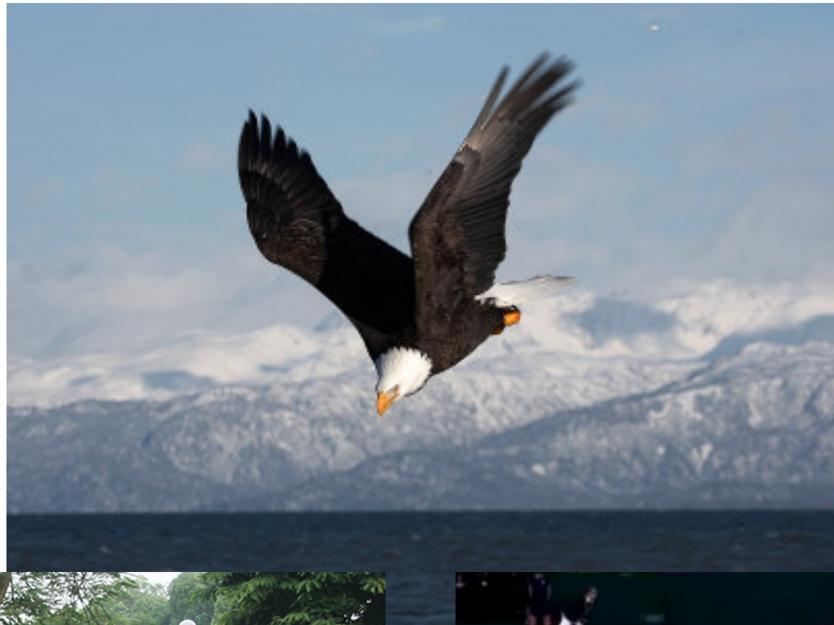
Normative solutions to the speed/accuracy trade-off in perceptual decision-making

Jan Drugowitsch
University of Geneva

Ruben Moreno-Bote
Alex Pouget
Mike Shadlen

FENS-Hertie Winter School 2015
The neuroscience of decision making

Perceptual decision-making?



Normative solutions?

Normative = how things ought to be

Identify
- the information on which decisions should be based
- the computations that need to be carried out on that information

Can be parameterised to fit the behaviour

Deviations from normative solutions give insight into specific deficits or biases

Resources

Notes, derivations, and figure code (MATLAB) is on

`https://github.com/jdrugo/FENS2015`

Road map

Normative evidence accumulation

What do we want to maximise?

Speed/accuracy trade-off for known reliability of evidence

Speed/accuracy trade-off for unknown reliability of evidence

Extensions

Road map

→ Normative evidence accumulation

What do we want to maximise?

Speed/accuracy trade-off for known reliability of evidence

Speed/accuracy trade-off for unknown reliability of evidence

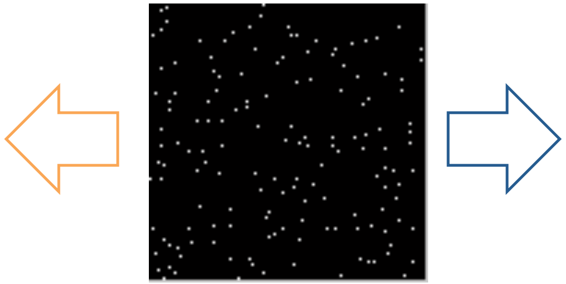
Extensions

Normative evidence accumulation

Random dot motion task

"left"? "right"?

51.2% coherence



12.8% coherence

Hidden state $z \in \{-1, 1\}$

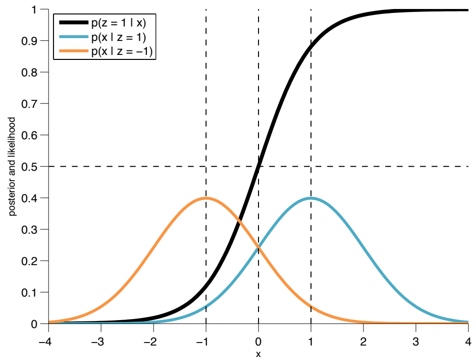
"left" "right"

Evidence $p(x \mid z) = \mathrm{N}\left(x \mid z, \sigma_\varepsilon^2\right)$

observation / stimulus

Posterior belief

$$p(z \mid x) \propto \overbrace{p(x \mid z)}^{\text{likelihood}} \overbrace{p(z)}^{\text{prior}} \propto \frac{1}{1 + e^{-2\frac{xz}{\sigma_\varepsilon^2}}}$$



`01evacc/single_obs.m`

Evidence accumulation with multiple observations

Hidden state $z \in \{-1,1\}$

"left" "right"

*n*th evidence $p(x_n \mid z) = \mathrm{N}\left(x_n \mid z, \sigma_\varepsilon^2\right)$

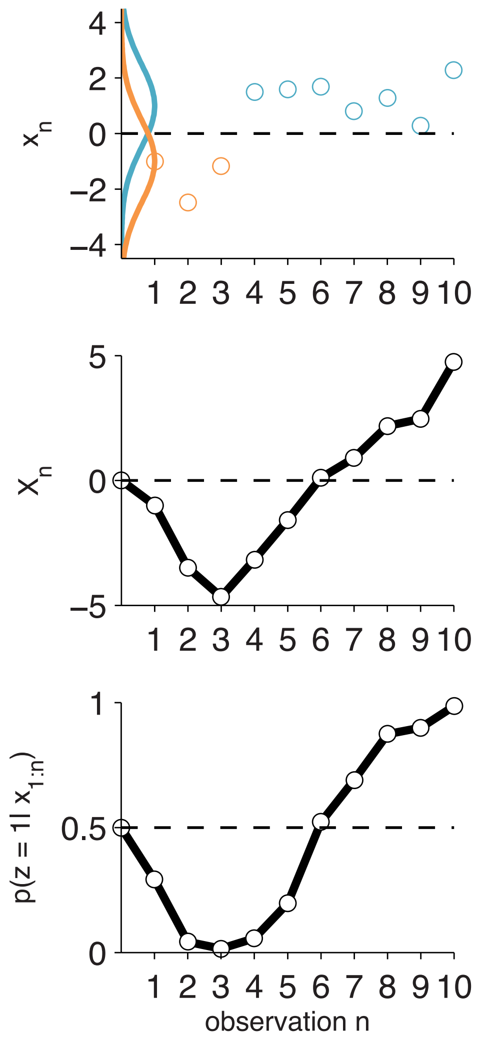
observation / stimulus

Posterior belief after *N* observations

prior i.i.d. likelihood

$$p(z \mid x_{1:N}) \propto p(z) \prod_{n=1}^{N} p(x_n \mid z) \propto \frac{1}{1+e^{-2\frac{X_N z}{\sigma_\varepsilon^2}}}$$

with *sufficient statistics* $X_N = \sum_{n=1}^{N} x_n$



01evacc/discrete_obs.m

Evidence accumulation with stream of evidence

Discretise stream into small chunks of δt

Likelihood of momentary evidence δx_n

$$p(\delta x_n \mid z) = \underbrace{\mathrm{N}\left(\delta x_n \mid z\delta t, \sigma_\varepsilon^2 \delta t\right)}$$

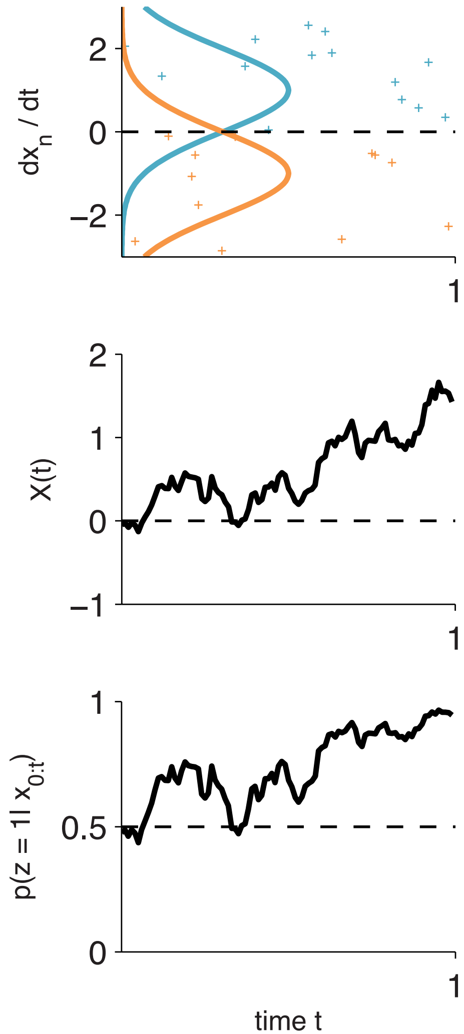
information in δx_n about z goes to 0 with $\delta t \rightarrow 0$

Posterior upon observing δx_1, δx_2, … *for t seconds (taking* $\delta t \rightarrow 0$)

$$p(z \mid \delta x_{0:t}) \propto \underbrace{p(z)}_{\text{prior}} \underbrace{\prod_{n=1}^{t/\delta t} p(\delta x_n \mid z)}_{\text{i.i.d. likelihood}} \propto \frac{1}{1 + e^{-2\frac{X(t)z}{\sigma_\varepsilon^2}}}$$

with *sufficient statistics* $X(t) = \underbrace{\int_0^t \delta x(s)}$

diffusion model $X(t) \sim \mathrm{N}\left(zt, \sigma_\varepsilon^2 t\right)$



01evacc/continuous_obs.m

Road map

Normative evidence accumulation

→ What do we want to maximise?

Speed/accuracy trade-off for known reliability of evidence

Speed/accuracy trade-off for unknown reliability of evidence

Extensions

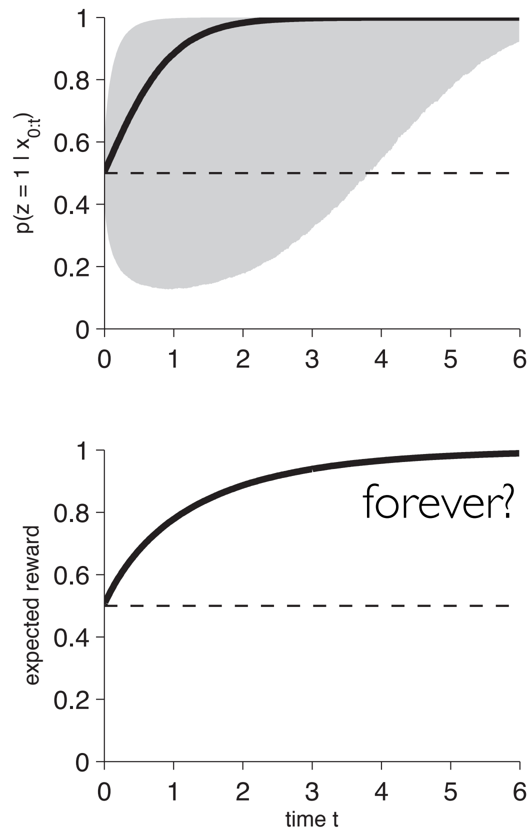
Maximizing 0-1 reward

To get from posterior to choices: need loss/reward function

reward(choice *a*, hidden state *z*)

Perform choice that maximises expected reward

how long to accumulate?



02lossfn/loss_01.m

0-1 reward

		hidden state	
		z=-1	z=1
choice	a="left"	1	0
	a="right"	0	1

→ choose more likely correct option

$p\left(z=1 \mid X(t)\right) \geq \frac{1}{2}$ → choose "right"

$p\left(z=1 \mid X(t)\right) < \frac{1}{2}$ → choose "left"

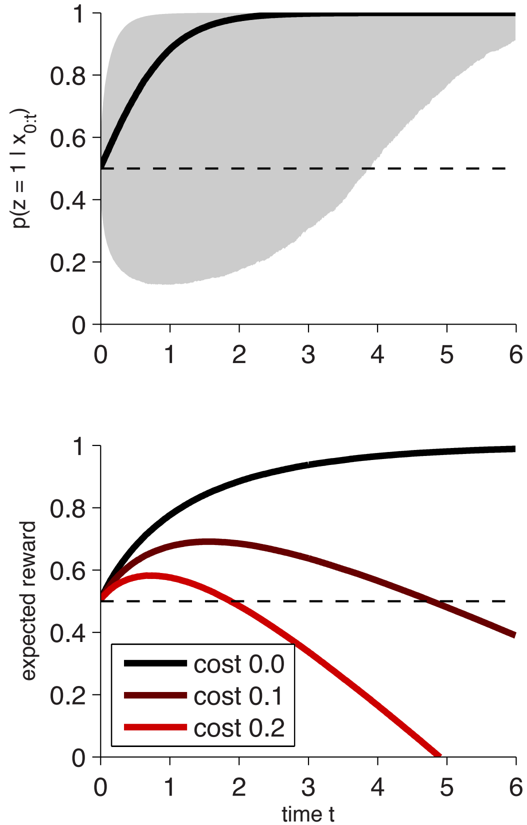
A cost for accumulating evidence

Cost *c* for accumulating evidence
- internal (e.g., effort / attention)
- external (e.g., lost future opportunities)

Expected reward

$$ER(PC, RT) = PC - cRT$$

PC ↙ p(correct) RT ↘ reaction time



After some time,
marginal choice accuracy increase does not justify associated cost

This is the reward function that we will use for the rest of the tutorial

`02lossfn/loss_cost.m`

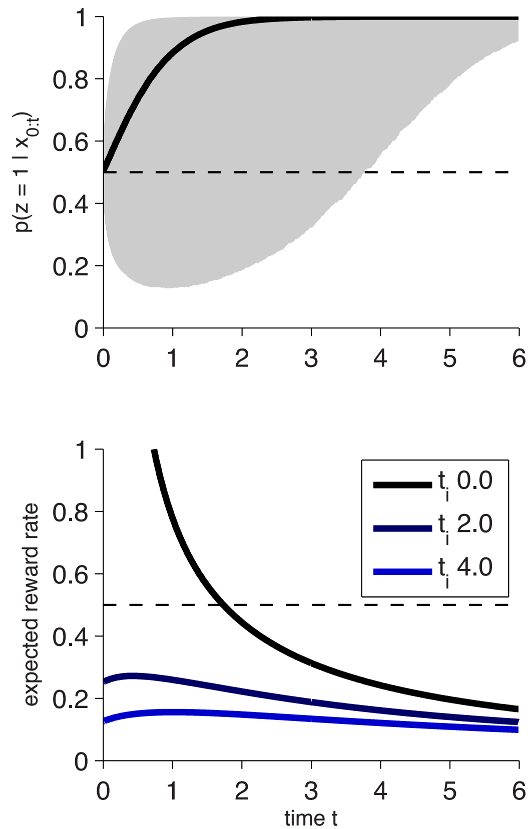
Maximising reward rate rather than expected reward

Loss of future reward is explicit in **reward rate**

$$RR(PC, RT) = \frac{PC - cRT}{RT + t_i + (1 - PC)t_p}$$

t_i: inter-trial interval

t_p: penalty time (for wrong choices)



02lossfn/loss_rr.m

Even c=0 leads to early choices,
as large *RT* reduces reward rate in denominator

We will come back to that as possible extension

Road map

Normative evidence accumulation

What do we want to maximise?

→ Speed/accuracy trade-off for known reliability of evidence

Speed/accuracy trade-off for unknown reliability of evidence

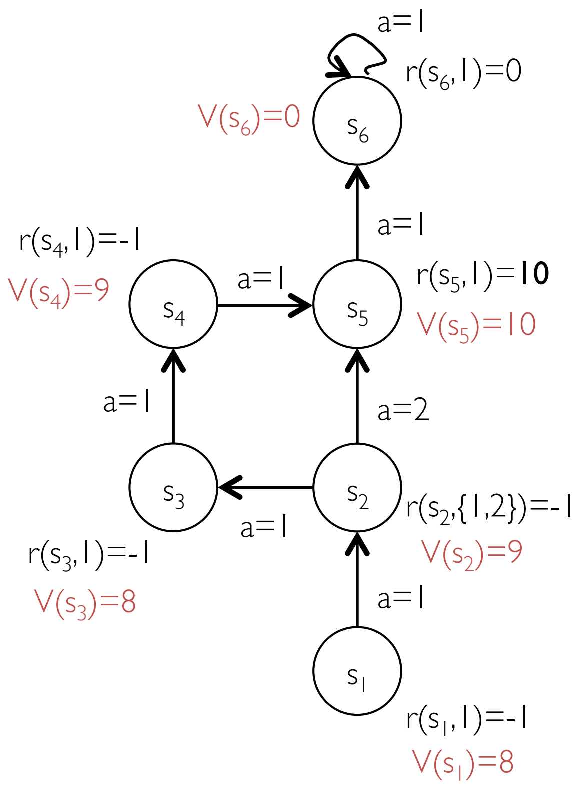
Extensions

Dynamic Programming in a nutshell

We find the optimal speed/accuracy trade-off by Dynamic Programming

Ingredients (Markov Decision Process):

- Set of states, $s \in S$
- Set of actions, $a \in A$
- State transition probabilities, $p(s' | s,a)$
- Reward function, $r(s,a)$
- (Discount factor, $\gamma = 1$)



Aim: find policy $\pi : S \rightarrow A$ that maximises total reward

$$V^{\pi}(s) = \left\langle \sum_{n=1}^{\infty} r\left(s_n, \pi(s_n)\right) \right\rangle_{p(s_1, s_2, \ldots | s_1 = s, \pi)}$$

from each state s.

Can be found recursively by solving **Bellman's equation:**

$$V(s) = \max_a \left[r(s,a) + \langle V(s') \rangle_{p(s'|s,a)} \right]$$

Optimal action (i.e., optimal policy) is action that maximises right-hand side.

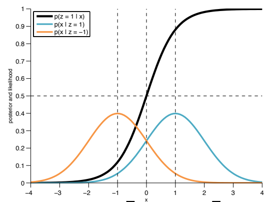
States and actions in perceptual decision-making

States: all observations so far → sufficient statistics $X \in [-\infty, \infty]$

To have bounded state-space, use *belief* $g \in [0,1]$ instead:

$$g(X) \equiv p(z=1 \mid X) = \frac{1}{1+e^{-2\frac{X}{\sigma_\varepsilon^2}}}$$

$g \in [0,1]$



$X \in [-\infty, \infty]$

`01evacc/single_obs.m`

Actions: $r(s,a) + \langle V(s') \rangle_{p(s'|s,a)}$

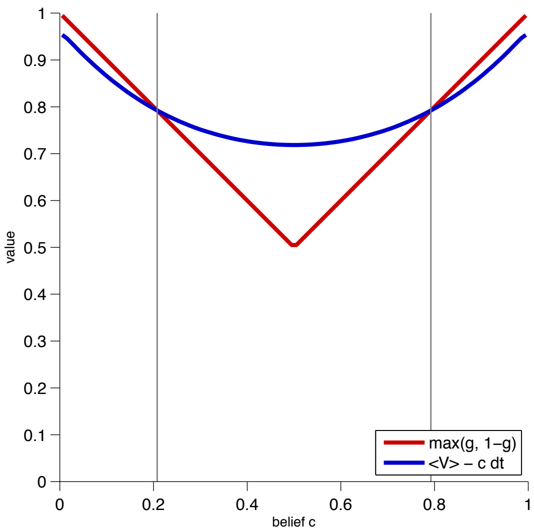
- choose "right" (*z*=1) $g + 0$
- choose "left" (*z*=-1) $1 - g + 0$
- accumulate another ***δt*** $-c\delta t + \langle V(g') \rangle_{p(g'|g)}$

Bellman's equation for perceptual decision-making:

$$V(g) = \max\left[g, 1-g, \langle V(g') \rangle_{p(g'|g)} - c\delta t\right]$$

→ **optimal decision-making by boundaries on belief *g***

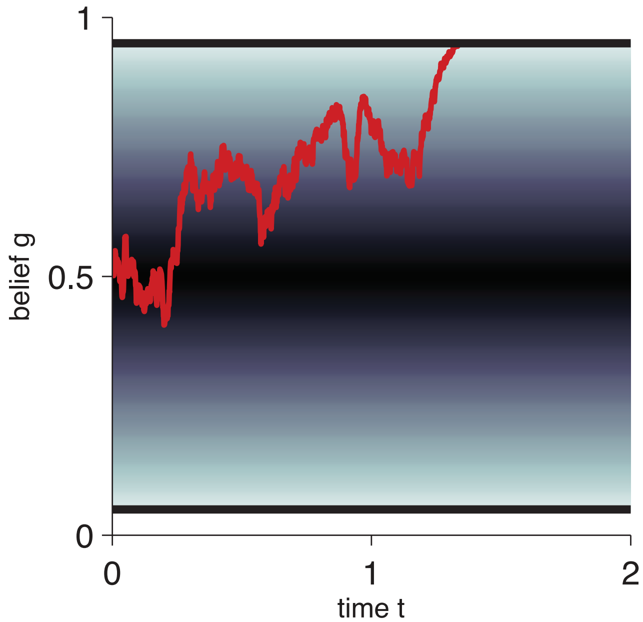
value



belief g

`03knownreliab/plot_valueintersect.m`

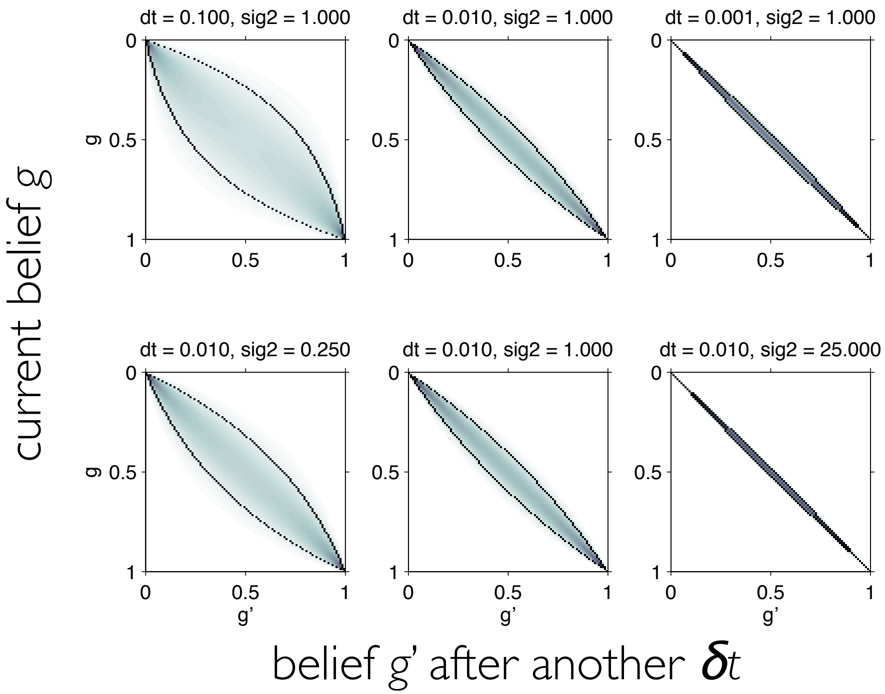
An example decision



03knownreliab/plot_diffusion_example.m

Finding the optimal bounds numerically

Find belief transition probability $p(g'|g)$ for accumulating δt more evidence



decreasing δt

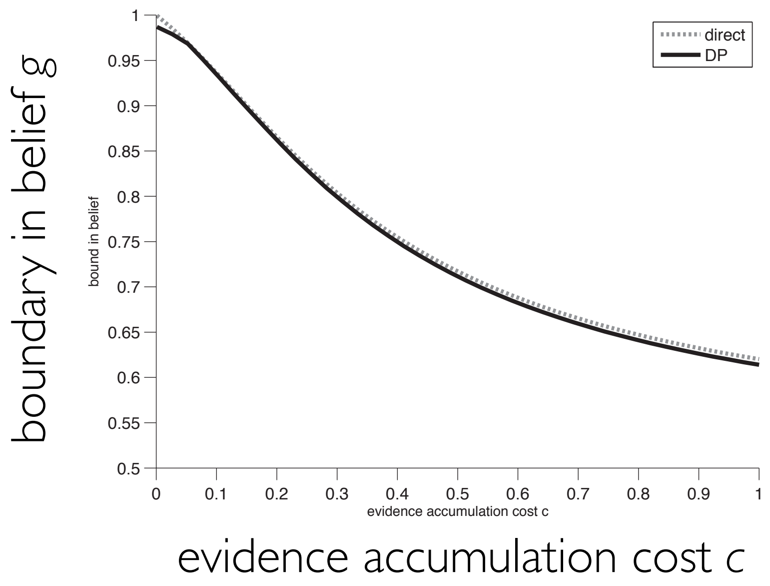
increasing σ_ε
(task difficulty)

`03knownreliab/plot_belieftrans.m`

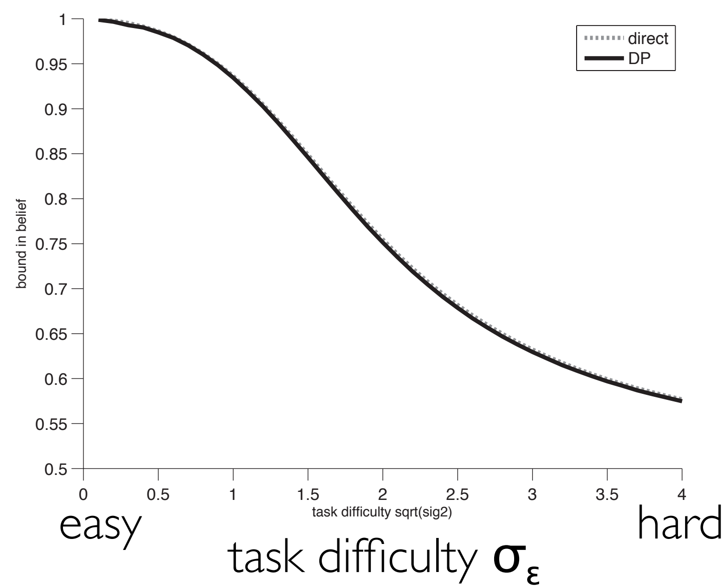
Discretise belief and value function into $k=1,\ldots K$, and solve numerically by value iteration:

$$V^{k,n} = \max\left[g^k, 1-g^k, \sum_{j=1}^{K} p\left(g^j \mid g^k\right) V^{j,n-1} - c\delta t\right]$$

Optimal boundary with cost, task difficulty



`03knownreliab/plot_bound_with_cost.m`



`03knownreliab/plot_bound_with_sig2.m`

An alternative to DP to find the optimal boundaries:

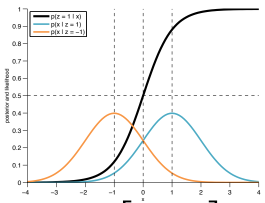
Maximising $ER(PC, RT) = PC - cRT$ using known expression for diffusion models

$$PC = \frac{1}{1 + e^{-2\frac{\theta}{\sigma_\varepsilon^2}}} \qquad RT = \theta \tanh \frac{\theta}{\sigma_\varepsilon^2}$$

Optimal decision-making with diffusion models

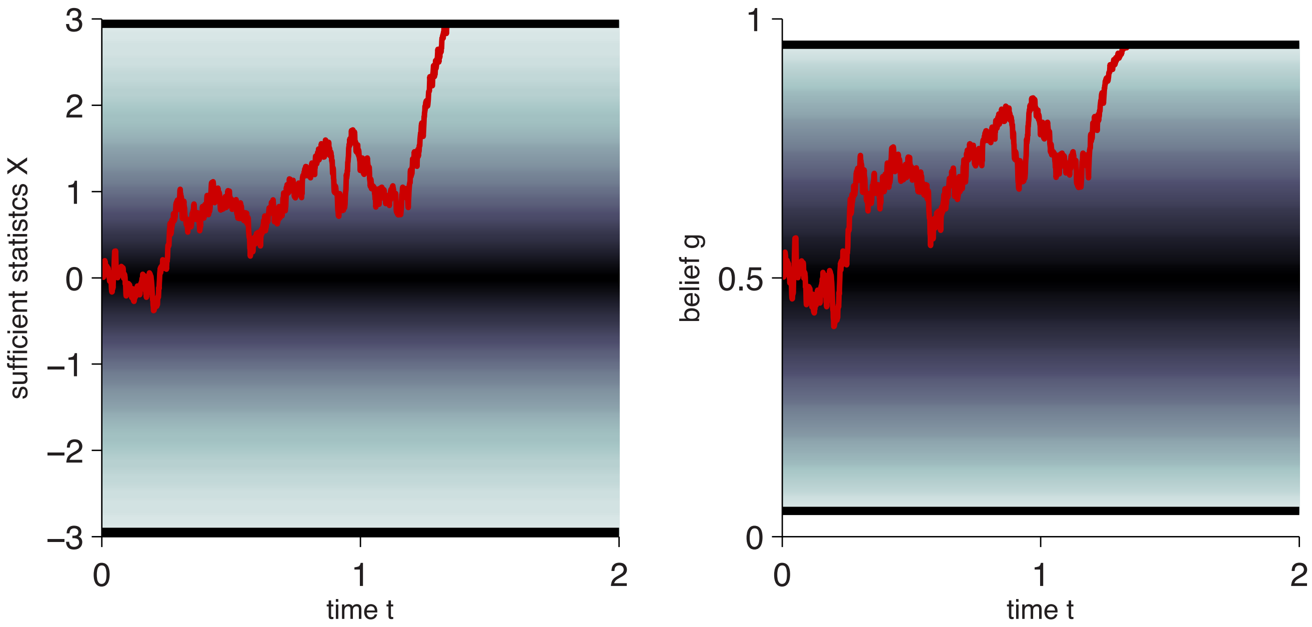
Boundaries on belief g → boundaries in X (diffusion model)

$g \in [0,1]$



$X \in [-\infty, \infty]$

01evacc/single_obs.m



03knownreliab/plot_diffusion_example.m

Can perform optimal decision-making without ever computing belief g

Road map

Normative evidence accumulation

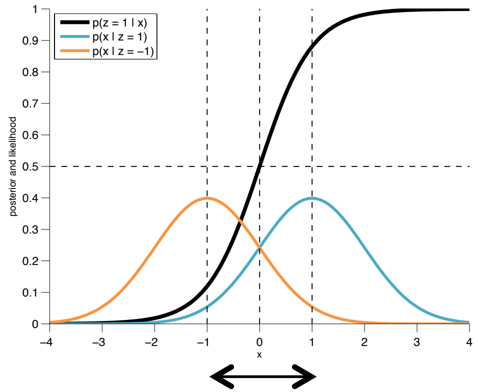
What do we want to maximise?

Speed/accuracy trade-off for known reliability of evidence

→ Speed/accuracy trade-off for unknown reliability of evidence

Extensions

Evidence accumulation with reliability changing across trials



defines trial difficulty for fixed $\sigma_\varepsilon = 1$

```
01evacc/single_obs.m
```

Hidden state, $z \in \{-1, 1\}$

Trial difficulty, $\alpha \in [0, \infty]$ (small = hard)

↓

Per-trial evidence mean $\mu = \alpha z$ drawn for each trial from

$$p(\mu) = \mathrm{N}\left(\mu \mid 0, \sigma_\alpha^2\right)$$

↘

overall task difficulty (small = hard)

Hard trials more likely, easy trials less likely

With evidence, $\delta x_n \sim \mathrm{N}\left(\mu \delta t, \delta t\right)$ evidence accumulation by

$$p\left(\mu \mid \delta x_{0:t}\right) \propto p(\mu) \prod_{n=1}^{t/\delta t} p\left(\delta x_n \mid \mu\right) \propto N\left(\mu \mid \frac{X(t)}{\sigma_\alpha^{-2} + t}, \frac{1}{\sigma_\alpha^{-2} + t}\right)$$

$$g(X,t) \equiv p\left(z = 1 \mid \delta x_{0:t}\right) = p\left(\mu \geq 0 \mid X(t), t\right) = \Phi\left(\frac{X(t)}{\sqrt{\sigma_\alpha^{-2} + t}}\right)$$

such that sufficient statistics are now both $X(t)$ and t

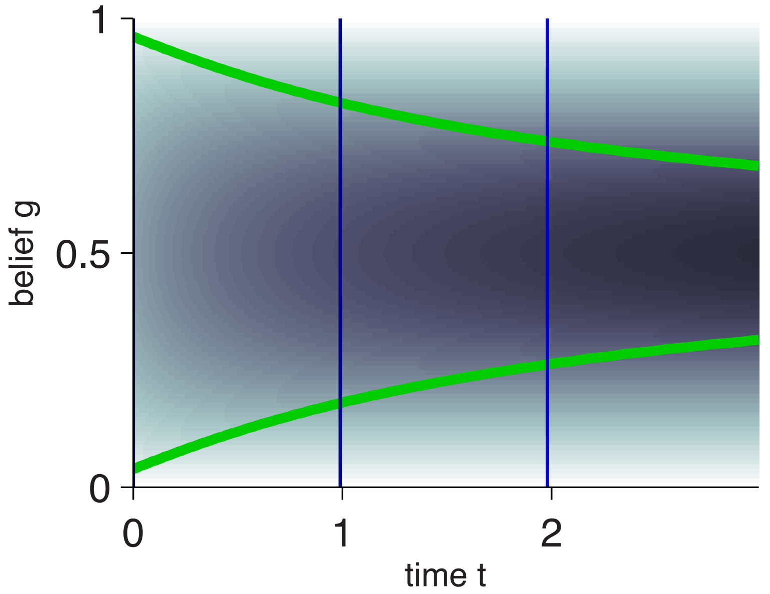
The resulting optimal policy

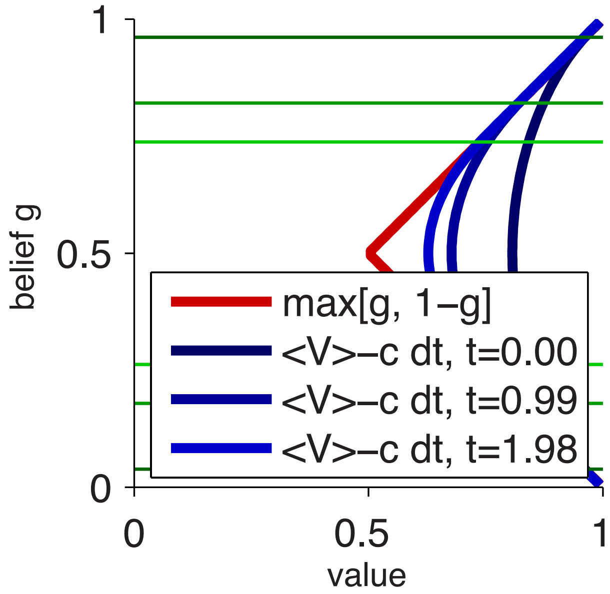
Bellman's equation over belief g and time t

$$V(g,t) = \max\left[g, 1-g, \underbrace{\left\langle V\left(g', t+\delta t\right)\right\rangle_{p(g'|g,t)}} - c\delta t\right]$$

future expected value takes into account time-dependence

For each fixed t → boundaries in belief g → time-dependent boundaries





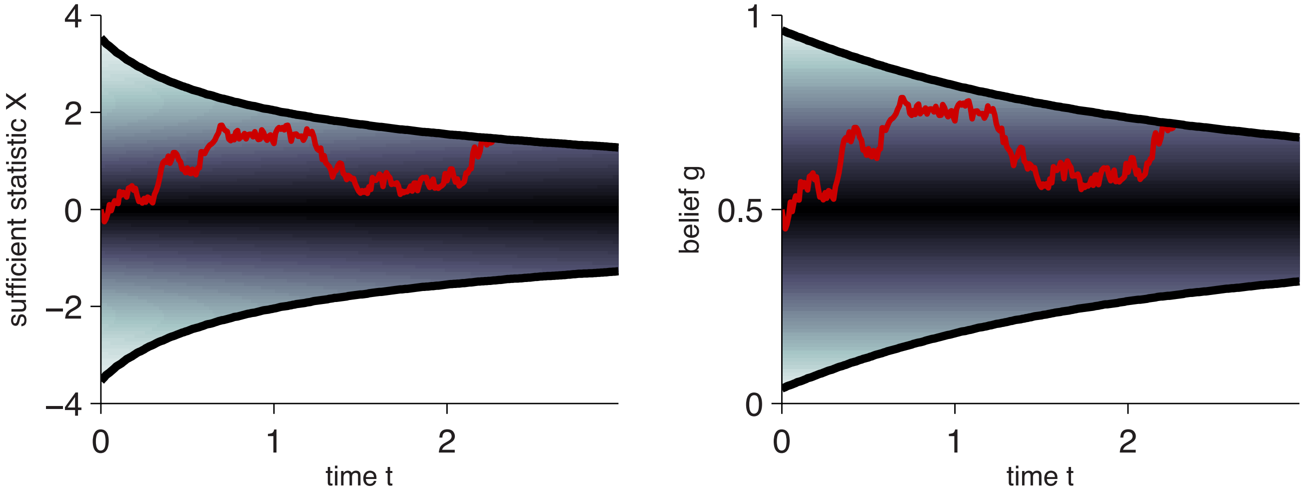
04knownreliab/plot_valuefn.m

Optimal decision-making with diffusion models

Using time-dependent mapping between belief g and diffusing particle X

$$g(X,t) = \Phi\left(\frac{X}{\sqrt{\sigma_\alpha^{-2}+t}}\right) \longleftrightarrow X(g,t) = \sqrt{\sigma_\alpha^{-2}+t}\,\Phi^{-1}(g)$$

Can map bound in belief g to bound in diffusing particle X

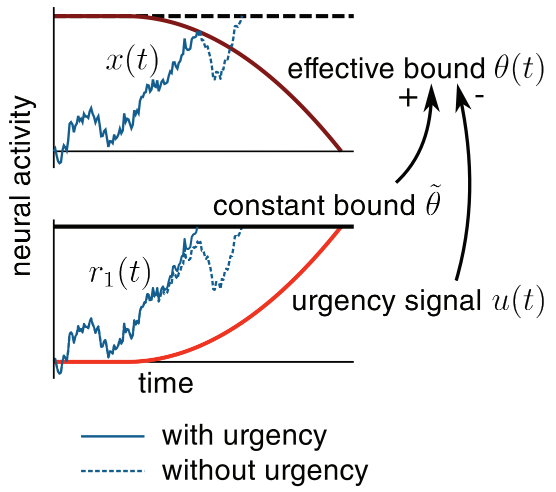


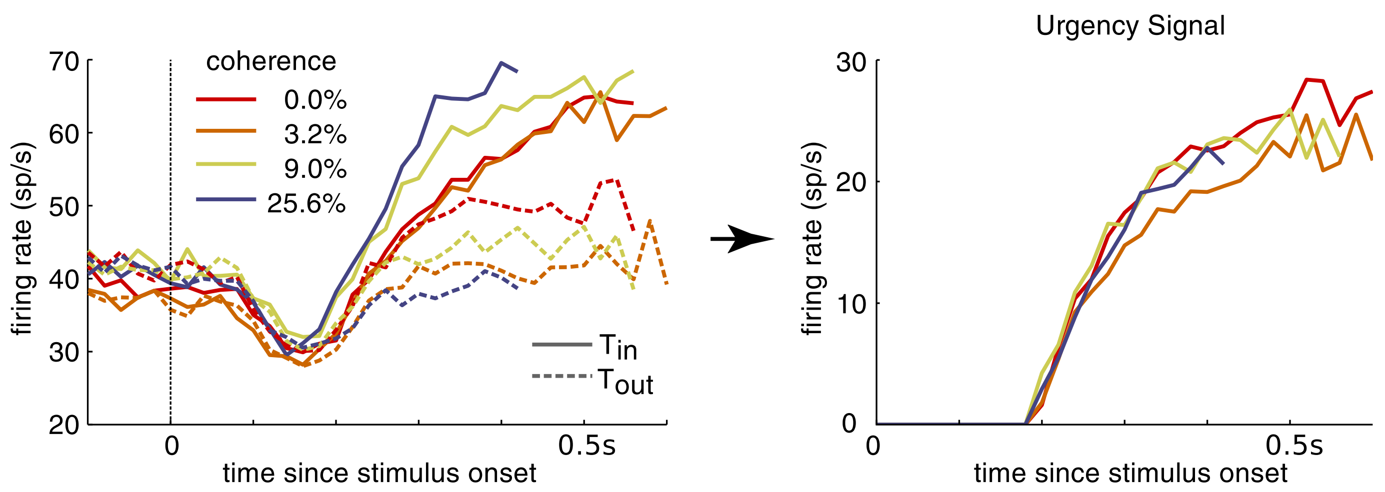
`04knownreliab/plot_diffusion_example.m`

As before, optimal decision-making without ever directly computing the belief
Naturally leads to slower error than correct choices

Urgency signal

"urgency signal" as collapsing bound?





Road map

Normative evidence accumulation

What do we want to maximise?

Speed/accuracy trade-off for known reliability of evidence

Speed/accuracy trade-off for unknown reliability of evidence

→ Extensions

Maximising reward rate instead of expected reward

Maximising reward rate = maximising reward in infinite sequence of structurally equal trials

Problem: value of initial choice will be infinite (infinite, possibly rewarding choices follow)

reward rate (reward per unit time)

Solution: use average-adjusted value, discounting time δt by $-\rho\delta t$

this value is relative to reward that can be achieved on average

For example, Bellman's equation for known evidence reliability

expected time from decision to onset of next trial

$$V(g) = \max\left[\begin{array}{c} g - \overbrace{\left(t_i + (1-g)t_p\right)}\rho + V\left(\tfrac{1}{2}\right), 1 - g - \overbrace{\left(t_i + g t_p\right)}\rho + V\left(\tfrac{1}{2}\right), \\ \left\langle V\left(g'\right)\right\rangle_{p(g'|g)} - (\rho + c)\delta t \end{array}\right]$$

Policy is invariant to shifts in value function, such that we can choose $V\left(\frac{1}{2}\right) = 0$.

Find both value function and reward rate, using consistency criterion, $V\left(\frac{1}{2}\right) = 0$.

Generalisations to evidence model

A time-varying accumulation cost

Accumulation cost might rise/drop over time

$$V(g,t) = \max\left[g, 1-g, \left\langle V\left(g', t+\delta t\right)\right\rangle_{p(g'|g,t)} - c(t)\delta t \right]$$

supported by human/animal behavior (Drugowitsch et al., 2012)

Evidence reliability that varies within individual trials

In real-world decisions, the evidence reliability is practically never constant.

Requires simultaneous estimation of hidden state and momentary evidence reliability
→ sufficient statistics are at least two-dimensional

Leads to reliability-dependent bound on decision-maker's belief,
see Drugowitsch, Moreno-Bote, Pouget (NIPS, 2014).

Paper also introduces faster way to find expected future reward, using PDE solvers.

Summary

Normative evidence accumulation

What do we want to maximise?

Speed/accuracy trade-off for known reliability of evidence

Speed/accuracy trade-off for unknown reliability of evidence

Extensions

Notes, derivations, and figure code (MATLAB) is on

`https://github.com/jdrugo/FENS2015`