Wilson 2014, Rational Obstinacy

Kevin Dorst — 24.223 Rationality

Idea: Memory loss, but not *total*: can pass on one of a limited set of messages (states of memory) to your future self.
→ Want to choose a protocol that makes the memory-states maximally informative.
→ When far more signals than memory, best protocol exhibits biases: strong (though not Kripkean) dogmatism, ignoring weak information, commutativity failures, influences of priors, and biased assimilation.

Effectively counts how many strong signals you've received, up to memory limit. Analogous to *weight-minded memory* from Singer et al.

The Model:
A state of the world, say B_H (coin is biased-heads) or B_T (biased-tails).

$B_H \rightsquigarrow P(H) = \frac{2}{3}$;
$B_T \rightsquigarrow P(H) = \frac{1}{3}$

A choice between a safe or risky action. Say, $a_S = 0$ and $a_R = \begin{cases} 1 \text{ if } B_H \\ -1 \text{ if } B_T \end{cases}$

A finite set of *memory states*, say $m_1, ..., m_5$.

A finite set of (independent) *signals* about the state, say s_0, s_1, s_2, s_3, where s_i = *tossed 3 times and got i heads*.

Likelihood ratios $\frac{P(s_i|B_H)}{P(s_i|B_T)} = \frac{1}{8}, \frac{1}{2}, \frac{2}{1}, \frac{8}{1}$.

A constant (small) risk of *termination*, η, at which point you take the action recommended by your plan given your memory state.

And, in Wilson's setting, go back to the initial state and keep playing forever

A Bayesian decision-maker with prior p_0 in B_H designs a *protocol* $\langle g^0, \sigma, d\rangle$ that includes

- A (distribution over) initial state(s) to start in g^0, or reset to;
- A *transition function* σ that, given a memory state and a signal, outputs a mixed strategy for shifting to a new memory state;
- A *decision rule d* which tells you which action to take given your state if the process terminates.

Intuitively, positive signals shifts up, negative ones down

Intuitively, take risky action iff high enough memory state to be confident of B_H, eg in m_4 or m_5.

The Bayesian DM designs a protocol that maximizes expected payoff per period, i.e. time between terminations.
Goal is to take risky option iff B_H, i.e. to wind up in a high enough state (say, m_4 or m_5) that you'll take the option if B_H, and to end up in a low enough state (say, m_1, m_2, m_3) that you'll decline it if B_T.

Given the true state (B_H or B_T)—which says how likely you are to get various signals—a choice of protocol induces a *Markov chain* on memory states.

Subject to the constraint that each state has a η chance of terminating (and so reverting to g^0).

For simplicity, suppose for a moment that $\eta = 0$.
Suppose ignore s_1 and s_2, always move up/down 1 when see s_0/ s_3.
Given B_H, $P(s_0) = (\frac{1}{3})^3 = \frac{1}{27}$, $P(s_1) = \frac{6}{27}$, $P(s_2) = \frac{12}{27}$, and $P(s_3) = \frac{8}{27}$.

So:

$$M_{B_H} = \begin{pmatrix} \frac{19}{27} & \frac{8}{27} & 0 & 0 & 0 \\ \frac{1}{27} & \frac{18}{27} & \frac{8}{27} & 0 & 0 \\ 0 & \frac{1}{27} & \frac{18}{27} & \frac{8}{27} & 0 \\ 0 & 0 & \frac{1}{27} & \frac{18}{27} & \frac{8}{27} \\ 0 & 0 & 0 & \frac{1}{27} & \frac{26}{27} \end{pmatrix}$$

And, given B_T, we get the 'inverted' chain:

$$M_{B_T} \begin{pmatrix} \frac{26}{27} & \frac{1}{27} & 0 & 0 & 0 \\ \frac{8}{27} & \frac{18}{27} & \frac{1}{27} & 0 & 0 \\ 0 & \frac{8}{27} & \frac{18}{27} & \frac{1}{27} & 0 \\ 0 & 0 & \frac{8}{27} & \frac{18}{27} & \frac{1}{27} \\ 0 & 0 & 0 & \frac{8}{27} & \frac{19}{27} \end{pmatrix}$$

Then we can ask: what are the long-run behaviors of these processes? How can we choose the transition-probabilities so optimize our chances of ending up in a high state iff B_H?

Time-slice reformulation: Can reformulate as an absent-minded-driver like situation, where each time-slice only knows m_i and the protocol chosen, so conditions on this (from the prior) and then transitions from their signal as they expect to be best.

This is the discussion of 'team equilibria'. Result is that optimal protocol is one that each time-slice wants to follow

Features

There are cutoff points in probabilities p_i[1] such that m_i effectively represents that range of credence.
→ The cutoff points say how high a signal needs to shift p_i in order to transition to a higher or lower memory state.
(It's as if you only have 5 different credal states!)

[1] Really, likelihood ratios $\frac{p_i}{1-p_i}$

I think it's that the prior, given that it finds itself in m_i, is in the range

I think...

As long as no signal is definitive, no state is absorbing.
→ No Kripkean dogmatism!

Don't want to let yourself get stuck

Now assume η is very low, so the process will run for a long time.

Then optimal protocol will *ignore non-extreme states*.

Save your memory states for *strong* signals, since you're confident those are coming

Will only move up/down memory states 1 at a time—effectively keeping track of how many strong signals on each side.

But not *just* doing this; some states are 'sticky': you don't always transition out of them when you get s_0 or s_3, but instead do so with some probability (adopting a mixed strategy).

Why? Intuitively, you'll be knocked out of m_5 too often. You'll be getting s_3 signals, but won't be able to move up, so when you get s_0 you'll drop out of it even though that s_0 signal probably has lots of evidence

In time-slice interpretation: if you wake up with m_5, you know the most recent signal was s_3, and that you have an abundance of s_3s; so when s_0 comes in, you think 'bah!'

against it which you've lost.
Need to choose exit probability so that, in expectation, will exit only when s_0s sufficiently numerous compared to s_3s.

How does reasoning actually work?

$\Rightarrow$ With high probability, the Markov chain will *mix*—reach it's steady-state, long-run probabilities—by the time the process terminates and a decision must be made.

Example of mixing:

- Consider whether it's sunny or rainy on a given day. If sunny, $\frac{3}{4}$ likely to be sunny the next day; if rainy, $\frac{1}{2}$ likely to be sunny the next day: $\begin{pmatrix} 3/4 & 1/4 \\ 1/2 & 1/2 \end{pmatrix}$
- Suppose sunny today. How likely to be sunny/rainy in *two* days?

$$\begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} 3/4 & 1/4 \\ 1/2 & 1/2 \end{pmatrix} \begin{pmatrix} 3/4 & 1/4 \\ 1/2 & 1/2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} 3/4 & 1/4 \\ 1/2 & 1/2 \end{pmatrix}^2 = \begin{pmatrix} \frac{11}{16} & \frac{5}{16} \end{pmatrix}$$

$\begin{pmatrix} 0.6875 & 0.3125 \end{pmatrix}$

- Generally, when we 'run' a Markov chain C for an arbitrarily long time n, then[2] the probability of being in any given state (the long-run proportion of time you spend in each state) is given by C^n.

[2] So long as all states are in one 'communicating class', so there are paths between each pair

$$\begin{pmatrix} 3/4 & 1/4 \\ 1/2 & 1/2 \end{pmatrix}^2 = \begin{pmatrix} 0.75 & 0.25 \\ 0.5 & 0.5 \end{pmatrix}^2 = \begin{pmatrix} 0.6875 & 0.3125 \\ 0.625 & 0.375 \end{pmatrix}$$
$$\begin{pmatrix} 0.75 & 0.25 \\ 0.5 & 0.5 \end{pmatrix}^3 \approx \begin{pmatrix} 0.672 & 0.328 \\ 0.656 & 0.344 \end{pmatrix}$$
$$\begin{pmatrix} 0.75 & 0.25 \\ 0.5 & 0.5 \end{pmatrix}^{10} \approx \begin{pmatrix} 0.666 & 0.333 \\ 0.666 & 0.333 \end{pmatrix}$$

- So in long-run, $\frac{2}{3}$ of days are sunny.

Back to 'why stickiness?'

Limiting distribution of $\begin{pmatrix} 0.0002 & 0.0017 & 0.014 & 0.109 & 0.875 \end{pmatrix}$
$\rightarrow$ Get to m_5 87.5% of the time, and to m_4 or m_5 (where correctly take risky action) 98.4% of the time.

And symmetrically for M_{B_T}

Suppose instead only exit m_1 or m_5 *half* the time when get strong opposing signal. Obstinate policy leads, given B_H to:

$$O_{B_H} = \begin{pmatrix} \frac{23}{27} & \frac{4}{27} & 0 & 0 & 0 \\ \frac{1}{27} & \frac{18}{27} & \frac{8}{27} & 0 & 0 \\ 0 & \frac{1}{27} & \frac{18}{27} & \frac{8}{27} & 0 \\ 0 & 0 & \frac{1}{27} & \frac{18}{27} & \frac{8}{27} \\ 0 & 0 & 0 & \frac{1}{54} & \frac{53}{54} \end{pmatrix}$$

Which has limiting distribution $\begin{pmatrix} 0.0002 & 0.00091 & 0.0073 & 0.0583 & 0.933 \end{pmatrix}$
$\rightarrow$ Gets to m_5 now 93.3% of the time and to m_4 or m_5 (where correctly take risky action) 99.2% of the time.

Optimal exit probability, given s_0, seems to be around $\frac{1}{50}$.

Though I just eyeballed it, rather than using her formulas; may be different if interior states become sticky, which they probably should

Rationalizing(?) biases:

- No Kripkean dogmatism, BUT:
- *Ignoring information.* Ignore all but most informative signals
- *Commutativity violations*
  $\langle s_3, s_3, s_0 \rangle$ will lead to probably staying in m_5, whereas
  $\langle s_3, s_0, s_3 \rangle$ will lead to being in m_4.
- *Confirmation bias:* if start with a higher prior for B_H, some interior states are *sticky down*: chance of going down, given s_0, is less than 1.
- *Biased assimilation:* If I have a high prior for B_H and you have a low prior, and we both receive $\langle s_0, s_3 \rangle$, then in expectation I'll go up (because of my interior sticky-down states) and you'll go up (because of your interior sticky-up states)

Again, when η is close to 0—unlike Dallmann, this form of obstinacy appears given *long run* evidence-gathering

And if low prior, some interior ones are *sticky up*

What should we make of all this?

1) If this models us, who is Bayesian DM?
2) Part of what's so interesting is that the biases are *stable*: you want to follow the (biased) protocol even once you're aware of them. (Team equilibrium interpretation.)
   Where does the following intuitive argument go wrong? 'I am really confident of B_H; but I know if I were less limited I would probably be less confident, so I should be less confident now'.
3) Worry: we do/can have pretty fine-grained credences! Although memory is limited in many ways, not obvious that it's limited in *this* way.

Unlike Kelly!

Does it hinge on not being able to become 'less' confident without dropping too far?