Credences Across Time: *Not the solutions!* Rather, just how Josh approached it....

Belief and its Limits (BaiL); Seminar 4 ANU
Instructor: Joshua Pearson; joshuaedwardpearson@gmail.com 28/10/25

Case 1 - Bias Detection

1 - Desiderata Katie is unsure whether the coin is fair or $\frac{3}{4}$ biased towards heads. So let's keep it simple and say: $C_0(F) = \frac{1}{2}$. She flips the coin and sees it land tails. This should raise her confidence that the coin is fair—tails is unlikely if the coin is biased towards heads: $C_1(F) > C_0(F)$.

2 - Evidence Katie sees the coin land tails. So plausibly her evidence at t_1, e_1, is something like *the coin landed tails.*

3 - Make-it-make-(Bayesian)-sense By conjunction, $C_0(F\&e_1) = C_0(e_1 \mid F)C_0(F)$. $C_0(e_1 \mid F)$ is Katie's credence the coin lands tails if it's fair, so $\frac{1}{2}$. Hence $C_0(F\&e_1) = \frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$. Similar calculations give us the following state space:

State	$C_0(\cdot)$
$F \wedge e_1$	$\frac{1}{4}$
$F \wedge \neg e_1$	$\frac{1}{4}$
$\neg F \wedge e_1$	$\frac{1}{8}$
$\neg F \wedge \neg e_1$	$\frac{3}{8}$

- Finally, let's check we meet the desiderata:
  - $C_0(F) = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$ ✓
  - $C_1(F) = \frac{C_0(F\&e_1)}{C_0(e_1)} = \frac{\frac{1}{4}}{\frac{1}{4}+\frac{1}{8}} = \frac{2}{3}$ ✓

Case 2 - Questionable Bias Detection

1 - Desiderata This is just Katie's case, except there's an additional time, t_2, where Al is told his coffee was spiked. Plausibly:

- $C_0(F) = \frac{1}{2}$
- $C_1(F) > C_0(F)$
- $C_2(F) \approx C_0(F)$

2 - Evidence Up to t_1, Al's case is just like Katie's, so at first glance it also seems like: e_1 =*the coin landed tails.* At t_2, Al then learns something like: *Josh says his coffee was spiked.*

Problem: this will not capture the desiderata for $C_2(F)$. Consider Al's initial conditional credence in F, given the coin landed tails and his coffee was spiked — $C_0(F \mid e_1 \& e_2)$. ***Given that the coin in fact landed tails, whether or not Al's coffee is spiked just seems completely irrelevant to whether the coin is fair.*** So, $C_0(F \mid e_1 \& e_2) = C_0(F \mid e_1) > C_0(Fair)$, contrary to our desiderata!

Solution: describe Al's evidence at t_1. Say e_1 isn't *The coin landed tails* but something more like *The coin appeared to land tails*.[1]

3 - **Make-it-make-(Bayesian)-sense** F and e_2 seem, at least initially, independent: $C_0(F \& e_2) = C_0(F) \& C_0(e_2)$. So, we can calculate the values for the state space using the **Conjunction** formula. For example:

$$C_0(F \& e_1 \& e_2) = C_0(e_1 \mid F \& e_2) C_0(F \& e_2) = C_0(e_1 \mid F \& e_2) C_0(F) C_0(e_2)$$

What is $C_0(e_2)$—Al's initial credence his coffee was spiked? The case doesn't say, but presumably it's low. Let's say $C_0(e_2) = \frac{1}{4}$.

So, for the first row, we just need to calculate $C_0(e_1 \mid F \& e_2)$—Al's initial credence the coin *appears* to land tails, given it's fair and his coffee was spiked. This is $\frac{1}{2}$—if his coffee is spiked Al is no better than chance at determining how a coin landed. So $C_0(F \& e_1 \& e_2) = (\frac{1}{2})(\frac{1}{2})(\frac{1}{4}) = \frac{1}{16}$.

For the second row we have:

$$C_0(F \& e_1 \& \neg e_2) = C_0(e_1 \mid F \& \neg e_2) C_0(F \& \neg e_2) = C_0(e_1 \mid F \& \neg e_2) C_0(F) C_0(\neg e_2)$$

What's $C_0(e_1 \mid F \& \neg e_2)$? This is Al's initial credence that the coin appears to land tails, given it's fair and his coffee *isn't* spiked. Let's keep it simple, and assume Al is perfectly reliable, in this scenario, at identifying how the coin lands. So $C_0(e_1 \mid F \& \neg e_2) = \frac{1}{2}$, and $C_0(e_1 \mid F \& \neg e_2) = (\frac{1}{2})(\frac{1}{2})(\frac{3}{4}) = \frac{3}{16}$.

State	C_0
$F \& e_1 \& e_2$	$C_0(e_1 \mid F \& e_2) C_0(F) C_0(e_2) = (\frac{1}{2})(\frac{1}{2})(\frac{1}{4}) = \frac{1}{16}$
$F \& e_1 \& \neg e_2$	$C_0(e_1 \mid F \& \neg e_2) C_0(F) C_0(\neg e_2) = (\frac{1}{2})(\frac{1}{2})(\frac{3}{4}) = \frac{3}{16}$
$F \& \neg e_1 \& e_2$	$C_0(\neg e_1 \mid F \& e_2) C_0(F) C_0(e_2) = (\frac{1}{2})(\frac{1}{2})(\frac{1}{4}) = \frac{1}{16}$
$F \& \neg e_1 \& \neg e_2$	$C_0(\neg e_1 \mid F \& \neg e_2) C_0(F) C_0(\neg e_2) = (\frac{1}{2})(\frac{1}{2})(\frac{3}{4}) = \frac{3}{16}$
$\neg F \& e_1 \& e_2$	$C_0(e_1 \mid \neg F \& e_2) C_0(\neg F) C_0(e_2) = (\frac{1}{2})(\frac{1}{2})(\frac{1}{4}) = \frac{1}{16}$
$\neg F \& e_1 \& \neg e_2$	$C_0(e_1 \mid \neg F \& \neg e_2) C_0(\neg F) C_0(\neg e_2) = (\frac{1}{4})(\frac{1}{2})(\frac{3}{4}) = \frac{3}{32}$
$\neg F \& \neg e_1 \& e_2$	$C_0(\neg e_1 \mid \neg F \& e_2) C_0(\neg F) C_0(e_2) = (\frac{1}{2})(\frac{1}{2})(\frac{1}{4}) = \frac{1}{16}$
$\neg F \& \neg e_1 \& \neg e_2$	$C_0(\neg e_1 \mid \neg F \& \neg e_2) C_0(\neg F) C_0(\neg e_2) = (\frac{3}{4})(\frac{1}{2})(\frac{3}{4}) = \frac{9}{32}$

- *Whew!* Now we can check we meet the desiderata...

[1] This is the "intrumentalist" approach Schoenfield describes—*Al must treat his perceptual inputs now more like reading an instrument, rather than as a direct source of evidence!* Note that there are many deep philosophical questions/problems here:

1. Does that mean my model of Katie, above, was *also* inaccurate? Should I have described her evidence as something more like *The coin appeared to land tails*? If so, is our evidence ever anything more than how things appear to us? (We often speak as if it can be! A lawyer who spoke purely in terms of the present constitution of their phenomenal experience wouldn't be particularly convincing...)
2. Suppose we considered a fourth case in which, at t_2, Al learns that he's terrible at identifying how things appear to him. That is, when he gets an appearance of tails, he'll often report the coin looks like it landed heads, and vice versa. Two questions:
   - Is such a case *even possible*?
   - If it is, does this mean Al's evidence such be *even weaker?* Something like: *Al has reported that the coin appeared to land on tails*? But then couldn't we argue that even *this* evidence should be weaker, so on and so forth?

- $C_0(F) = \frac{1}{16} + \frac{3}{16} + \frac{1}{16} + \frac{3}{16} = \frac{1}{2}$ ✓
- $C_1(F) = \frac{C_0(F \& e_1)}{C_0(e_1)} = \frac{\frac{1}{16}+\frac{3}{16}}{\frac{1}{16}+\frac{3}{16}+\frac{1}{16}+\frac{3}{32}} = \frac{\frac{4}{16}}{\frac{13}{32}} = \frac{8}{13}$ ✓[2]
- $C_2(F) = \frac{C_1(F \& e_2)}{C_1(e_2)} = \frac{\frac{1}{16}}{\frac{1}{16}+\frac{1}{16}} = \frac{1}{2}$ ✓

[2] That's a little less than Katie's confidence of F at t_1; this seems right given we've suggested Al's evidence is weaker than what we assumed for Katie's evidence.

Case 3 - Hypoxia

Cases like Hypoxia have been the topic of numerous papers in formal epistemology for the past 20-ish years. Any model of it will be controversial in some respect...

1 - Desiderata Plausibly:

- $C_0(L) = \textit{middling}$; let's just say $\frac{1}{2}$.
- $C_1(L) > C_0(L)$
- $C_2(L) \approx C_0(L)$

2 - Evidence The case says that, at t_1, Miriam "looks at her coordinates, her fuel tank gauge, her aircraft's weight, etc—any relevant information available to her." It's tempting to make e_1 the content of the information available to her. But in line with Case 2, let's say e_2 is just *the relevant information appears to indicate I*. And e_2 is *Ground control claim Miriam is Hypoxic*.

Problem. This is not going to capture our desiderata! The problem is similar to the one we encountered with Case 2.

Note that, first, we assume $C_1(L) = C_0(L \mid e_1) > C_0(L)$. That is, Miriam's evidence—*the relevant information appears to indicate I*—supports L.

Given this, what's $C_0(L \mid e_1 \& e_2)$? That is, what is Miriam's initial credence that she'll have enough fuel to get to LA, conditional on the information appearing to indicate *I and* that ground control claim she's hypoxic at t_1?

Remember: **hypoxia does not make Miriam bad at *identifying* evidence** (unlike the drug I supposedly spiked Al's coffee with!), **it just makes her bad at *assessing* that evidence.**

So, it seems like, given her evidence, and given that her evidence in fact supports L, ***whether Miriam has hypoxia or not is just irrelevant as to whether*** L. So $C_0(L \mid e_1 \& e_2) = C_0(L \mid e_1) > C_0(L)$, contrary to our desiderata![3]

[3] Some people just consider this a *result*: Miriam should not lower her credence in L at t_2. I can see how someone gripped within the Bayesian framework might reason themselves to this conclusion.... but it is an odd one, isn't it!?

Solution? Perhaps Miriam isn't treating herself like an instrument sufficiently enough. Not only should she treat her ability to identify information like a thermometer, she should also treat her ability to *assess* information like a thermometer, too. This would make Miriam's the better candidate for e_1 something more like: *Miriam's on board assessment is that her evidence supports L.*

3 - Make-it-make-(Bayesian)-sense With this apparent solution, we can appear to make good sense of the case. For simplicity, assume $C_0(e_2) = \frac{1}{4}$, and note again that L and e_2 are plausibly independent: $C_0(L\&e_2) = C_0(L)C_0(e_2)$. We now just need to decide how good Miriam is at assessing evidence when she isn't hypoxic, and how often the evidence supports the truth. Assume she's perfectly reliable without hypoxia and that the evidence is never misleading.[4]

[4] This is just for simplicity; capturing the desiderata doesn't depend on these two choices.

State	C_0
$L\&e_1\&e_2$	$C_0(e_1 \mid L\&e_2)C_0(L)C_0(e_2) = (\frac{1}{2})(\frac{1}{2})(\frac{1}{4}) = \frac{1}{16}$
$L\&e_1\&\neg e_2$	$C_0(e_1 \mid L\&\neg e_2)C_0(L)C_0(\neg e_2) = (1)(\frac{1}{2})(\frac{3}{4}) = \frac{3}{8}$
$L\&\neg e_1\&e_2$	$C_0(\neg e_1 \mid L\&e_2)C_0(L)C_0(e_2) = (\frac{1}{2})(\frac{1}{2})(\frac{1}{4}) = \frac{1}{16}$
$L\&\neg e_1\&\neg e_2$	$C_0(\neg e_1 \mid L\&\neg e_2)C_0(L)C_0(\neg e_2) = (0)(\frac{1}{2})(\frac{3}{4}) = 0$
$\neg L\&e_1\&e_2$	$C_0(e_1 \mid \neg L\&e_2)C_0(\neg L)C_0(e_2) = (\frac{1}{2})(\frac{1}{2})(\frac{1}{4}) = \frac{1}{16}$
$\neg L\&e_1\&\neg e_2$	$C_0(e_1 \mid \neg L\&\neg e_2)C_0(\neg L)C_0(\neg e_2) = (0)(\frac{1}{2})(\frac{3}{4}) = 0$
$\neg L\&\neg e_1\&e_2$	$C_0(\neg e_1 \mid \neg L\&e_2)C_0(\neg L)C_0(e_2) = (\frac{1}{2})(\frac{1}{2})(\frac{1}{4}) = \frac{1}{16}$
$\neg L\&\neg e_1\&\neg e_2$	$C_0(\neg e_1 \mid \neg L\&\neg e_2)C_0(\neg L)C_0(\neg e_2) = (1)(\frac{1}{2})(\frac{3}{4}) = \frac{3}{8}$

- $C_0(L) = \frac{1}{16} + \frac{1}{16} + \frac{3}{8} = \frac{1}{2}$ ✓
- $C_1(L) = \frac{C_0(L\&e_1)}{C_0(e_1)} = \frac{\frac{1}{16}+\frac{3}{8}}{\frac{1}{16}+\frac{3}{8}+\frac{1}{16}} = \frac{\frac{7}{16}}{\frac{8}{16}} = \frac{7}{8}$ ✓
- $C_2(L) = \frac{C_1(L\&e_2)}{C_1(e_2)} = \frac{\frac{1}{16}}{\frac{1}{16}+\frac{1}{16}} = \frac{1}{2}$ ✓

So... why is Schoenfield dissatisfied with this? Isn't this a perfectly good model of the case??

Here's my interpretation of Schoenfield's complaint. Consider our specified e_1 again:

e_1=*Miriam's on board assessment is that her evidence supports L.*

What is "her evidence" referring to here?[5] And does this evidence of hers, on it's own, support L or not?

[5] It can't be referring to e_1—Miriam isn't assessing her own assessment.

Answering this question puts us in a dilemma:

- If "her evidence" supports L, then it does so irrespective of whether she's hypoxic—and so $C_2(L) = C_1(L) > C_0(L)$.[6]
- If "her evidence" does *not* support L, then when Miriam is not hypoxic, she should *not* become confident in L—she'll perfectly assess that her evidence is no support of L! So $C_1(L) = C_0(L)$
- Either way, it's impossible to meet the desiderata!

[6] This is essentially the same problem as when we assumed e_1 was *Her relevant information appears to indicate I.*

So... how should we model the case? I have no idea![7]

[7] But see the end of Schoenfield's paper for an interesting suggestion.