Douven's Model *vs* Our Bayesian Model of "Abductive Updating"

Branden Fitelson
August 2, 2022

In[∘]:=
```
<< PrSAT`
```

Here is a Bayesian model of an agent's priors $\mathbf{cr}_{t_0}(\bullet)$, which are such that:

(1) $\mathbf{cr}_{t_0}(\bullet)$ is regular,

(2) $\mathbb{E}$ *confirms* $\mathbb{H}$: $\mathbf{cr}_{t_0}(\mathbb{H} \mid \mathbb{E}) > \mathbf{cr}_{t_0}(H \mid {\sim}\mathbb{E})$,

(3) **ERA** holds: $\mathbf{cr}_{t_0}(\mathbb{H} \mid \mathbb{E} \,\&\, \mathbb{A}) > \mathbf{cr}_{t_0}(\mathbb{H} \mid \mathbb{E} \,\&\, {\sim}\mathbb{A})$,

- Here, $\mathbb{A}$ is interpreted as asserting that $\mathbb{H}$ is the best (or only) explanation of $\mathbb{E}$.

(4) $\mathbb{E}$ *is (a priori) irrelevant to* $\mathbb{A}$: $\mathbf{cr}_{t_0}(\mathbb{A} \mid \mathbb{E}) = \mathbf{cr}_{t_0}(\mathbb{A} \mid {\sim}\mathbb{E})$,

(5) $\mathbb{A}$ *is (a priori) irrelevant to* $\mathbb{H}$: $\mathbf{cr}_{t_0}(\mathbb{H} \mid \mathbb{A}) = \mathbf{cr}_{t_0}(\mathbb{H} \mid {\sim}\mathbb{A})$, and

(6) $\mathbf{cr}_{t_0}(\mathbb{H}) = \mathbf{cr}_{t_0}(\mathbb{E}) = \mathbf{cr}_{t_0}(\mathbb{A}) = \frac{1}{2}$.

In[∘]:=
```
prior = PrSAT[{
    Pr[ℍ | 𝔼] > Pr[ℍ | ¬ 𝔼],
    Pr[ℍ | 𝔼 ∧ 𝔸] > Pr[ℍ | 𝔼 ∧ ¬ 𝔸],
    Pr[𝔸 | 𝔼] == Pr[𝔸 | ¬ 𝔼],
    Pr[ℍ | 𝔸] == Pr[ℍ | ¬ 𝔸],
    Pr[ℍ] == 1/2, Pr[𝔼] == 1/2, Pr[𝔸] == 1/2
   }, Probabilities → Regular, BypassSearch → True]
```

Out[∘]=

$$\left\{\{\mathbb{A} \to \{a_2, a_5, a_6, a_8\}, \mathbb{E} \to \{a_3, a_5, a_7, a_8\}, \mathbb{H} \to \{a_4, a_6, a_7, a_8\}, \Omega \to \{a_1, a_2, a_3, a_4, a_5, a_6, a_7, a_8\}\},\right.$$
$$\left.\left\{a_1 \to \frac{3}{32}, a_2 \to \frac{7}{32}, a_3 \to \frac{5}{32}, a_4 \to \frac{5}{32}, a_5 \to \frac{1}{32}, a_6 \to \frac{1}{32}, a_7 \to \frac{3}{32}, a_8 \to \frac{7}{32}\right\}\right\}$$

𝔸	𝔼	ℍ	cr_{t_0}
T	T	T	$\frac{7}{32}$
T	T	F	$\frac{1}{32}$
T	F	T	$\frac{1}{32}$
T	F	F	$\frac{7}{32}$
F	T	T	$\frac{3}{32}$
F	T	F	$\frac{5}{32}$
F	F	T	$\frac{5}{32}$
F	F	F	$\frac{3}{32}$

If we look at this agent's credences over the sub-algebra {ℍ,𝔼}, we have the following (coarse-graining of their) prior.

In[ ]:=
```
EvaluateProbability[{Pr[𝔼 ∧ ℍ], Pr[𝔼 ∧ ¬ ℍ], Pr[¬ 𝔼 ∧ ℍ], Pr[¬ 𝔼 ∧ ¬ ℍ]}, prior]
```

Out[ ]=

$$\left\{\frac{5}{16}, \frac{3}{16}, \frac{3}{16}, \frac{5}{16}\right\}$$

𝔼	ℍ	cr_{t_0}
T	T	$\frac{5}{16}$
T	F	$\frac{3}{16}$
F	T	$\frac{3}{16}$
F	F	$\frac{5}{16}$

Here's what the Bayesian evolution of this prior looks like, first at t_0 (before anything is learned), then at t_1 (when 𝔼 is learned), and finally at t_2 (when 𝔸 is learned).

In[ ]:=
```
EvaluateProbability[{Pr[𝔸 ∧ 𝔼 ∧ ℍ], Pr[𝔸 ∧ 𝔼 ∧ ¬ ℍ], Pr[𝔸 ∧ ¬ 𝔼 ∧ ℍ], Pr[𝔸 ∧ ¬ 𝔼 ∧ ¬ ℍ],
  Pr[¬ 𝔸 ∧ 𝔼 ∧ ℍ], Pr[¬ 𝔸 ∧ 𝔼 ∧ ¬ ℍ], Pr[¬ 𝔸 ∧ ¬ 𝔼 ∧ ℍ], Pr[¬ 𝔸 ∧ ¬ 𝔼 ∧ ¬ ℍ]}, prior]
```

Out[ ]=

$$\left\{\frac{7}{32}, \frac{1}{32}, \frac{1}{32}, \frac{7}{32}, \frac{3}{32}, \frac{5}{32}, \frac{5}{32}, \frac{3}{32}\right\}$$

In[ ]:=
```
EvaluateProbability[{Pr[𝔸 ∧ 𝔼 ∧ ℍ | 𝔼], Pr[𝔸 ∧ 𝔼 ∧ ¬ ℍ | 𝔼],
  Pr[𝔸 ∧ ¬ 𝔼 ∧ ℍ | 𝔼], Pr[𝔸 ∧ ¬ 𝔼 ∧ ¬ ℍ | 𝔼], Pr[¬ 𝔸 ∧ 𝔼 ∧ ℍ | 𝔼],
  Pr[¬ 𝔸 ∧ 𝔼 ∧ ¬ ℍ | 𝔼], Pr[¬ 𝔸 ∧ ¬ 𝔼 ∧ ℍ | 𝔼], Pr[¬ 𝔸 ∧ ¬ 𝔼 ∧ ¬ ℍ | 𝔼]}, prior]
```

Out[ ]=

$$\left\{\frac{7}{16}, \frac{1}{16}, 0, 0, \frac{3}{16}, \frac{5}{16}, 0, 0\right\}$$

In[∘]:=
```
EvaluateProbability[{Pr[𝔸 ∧ 𝔼 ∧ ℍ | 𝔼 ∧ 𝔸], Pr[𝔸 ∧ 𝔼 ∧ ¬ ℍ | 𝔼 ∧ 𝔸],
  Pr[𝔸 ∧ ¬ 𝔼 ∧ ℍ | 𝔼 ∧ 𝔸], Pr[𝔸 ∧ ¬ 𝔼 ∧ ¬ ℍ | 𝔼 ∧ 𝔸], Pr[¬ 𝔸 ∧ 𝔼 ∧ ℍ | 𝔼 ∧ 𝔸],
  Pr[¬ 𝔸 ∧ 𝔼 ∧ ¬ ℍ | 𝔼 ∧ 𝔸], Pr[¬ 𝔸 ∧ ¬ 𝔼 ∧ ℍ | 𝔼 ∧ 𝔸], Pr[¬ 𝔸 ∧ ¬ 𝔼 ∧ ¬ ℍ | 𝔼 ∧ 𝔸]}, prior]
```
Out[∘]=

$\left\{\frac{7}{8}, \frac{1}{8}, 0, 0, 0, 0, 0, 0\right\}$

𝔸	𝔼	ℍ	cr_{t_0}	cr_{t_1}	cr_{t_2}
T	T	T	$\frac{7}{32}$	$\frac{7}{16}$	$\frac{7}{8}$
T	T	F	$\frac{1}{32}$	$\frac{1}{16}$	$\frac{1}{8}$
T	F	T	$\frac{1}{32}$	0	0
T	F	F	$\frac{7}{32}$	0	0
F	T	T	$\frac{3}{32}$	$\frac{3}{16}$	0
F	T	F	$\frac{5}{32}$	$\frac{5}{16}$	0
F	F	T	$\frac{5}{32}$	0	0
F	F	F	$\frac{3}{32}$	0	0

What does Douven’s model of an analogous agent look like? Well, presumably, it starts off the same, but only on the sub-algebra {ℍ,𝔼}, since the agent is not assumed to have 𝔸 in their algebra of entertained propositions.

𝔼	ℍ	cr_{t_0}
T	T	$\frac{5}{16}$
T	F	$\frac{3}{16}$
F	T	$\frac{3}{16}$
F	F	$\frac{5}{16}$

Here, if the agent were to follow Bayes’s Theorem, then they would assign a value $\frac{5}{8}$ to cr_{t_0}(ℍ | 𝔼).

In[∘]:=
```
EvaluateProbability[Pr[ℍ | 𝔼], prior]
```
Out[∘]=

$\frac{5}{8}$

Moreover, if we look at the t_2 Bayesian posterior — after learning both 𝔼 and 𝔸, *viz.*, cr_{t_0}(ℍ | 𝔼 ∧ 𝔸) — we have the value $\frac{7}{8}$.

In[∘]:=
```
EvaluateProbability[Pr[ℍ | 𝔼 ∧ 𝔸], prior]
```
Out[∘]=

$\frac{7}{8}$

On the other hand, if the agent follows Douven’s rule for updating on 𝔼 — assuming that 𝔸 is true, *i.e.*, assuming 𝔠(ℍ,𝔼) > 0 — then we will have:

$$\mathbf{cr}_{t_0}(\mathbb{H} \mid \mathbb{E}) = \frac{cr_{t_0}(\mathbb{H})\, cr_{t_0}(\mathbb{E} \mid \mathbb{H}) + \mathfrak{c}(\mathbb{H},\mathbb{E})}{cr_{t_0}(\mathbb{H})\, cr_{t_0}(\mathbb{E} \mid \mathbb{H}) + \mathfrak{c}(\mathbb{H},\mathbb{E}) + cr_{t_0}(\neg\mathbb{H})\, cr_{t_0}(\mathbb{E} \mid \neg\mathbb{H})}$$

In[ ]:=
```
DouvenPosterior[c_, pr_] :=
 EvaluateProbability[(Pr[H] × Pr[E | H] + c) / (Pr[H] × Pr[E | H] + c + Pr[¬ H] × Pr[E | ¬ H]), pr]
```

For instance, if $\mathfrak{c}(\mathbb{H},\mathbb{E}) = \frac{1}{10}$, then we will obtain the value $\frac{11}{16}$ for the (Douven) posterior.

In[ ]:=
```
DouvenPosterior[1/10, prior]
```

Out[ ]=

$\frac{11}{16}$

Interestingly, ***there is no value of*** $\mathfrak{c}(\mathbb{H},\mathbb{E}) \in (0,1)$ such that the Bayesian $\mathbf{t}_2$ posterior $\mathbf{cr}_{t_0}(\mathbb{H} \mid \mathbb{E} \wedge \mathbb{A})$ *is equal to* the Douven posterior $\mathbf{cr}_{t_0}(\mathbb{H} \mid \mathbb{E})$! Specifically, ***only if*** $\mathfrak{c}(\mathbb{H},\mathbb{E}) = 1$ will the two models will agree on the $\mathbf{t}_2$ posterior.

In[ ]:=
```
Reduce[DouvenPosterior[c, prior] == 7/8, c]
```

Out[ ]=

```
c == 1
```

In[ ]:=
```
DouvenPosterior[1, prior]
```

Out[ ]=

$\frac{7}{8}$

Ultimately, these two agents will "look the same" at t2 — over the {ℍ,𝔼} sub-algebra. To wit:

𝔼	ℍ	$\mathbf{cr}_{t_2}$
T	T	$\frac{7}{8}$
T	F	$\frac{1}{8}$
F	T	0
F	F	0

The advantage of the Bayesian approach is that — assuming probabilism — the prior $\mathbf{cr}_{t_0}$ *completely determines* what happens in subsequent updates (*e.g.,* at $\mathbf{cr}_{t_1}$ and $\mathbf{cr}_{t_2}$). There is no "extra parameter" $\mathfrak{c}(\mathbb{H},\mathbb{E})$ that has to be fixed in order to determine $\mathbf{cr}_{t_2}$. Also, we don't need any "exogenous perturbation" to the agent's credences — it's just standard conditionalization (first on 𝔼 and then on 𝔸).

Moreover, I would suggest that the best way to fix the value of $\mathfrak{c}(\mathbb{H},\mathbb{E})$ is by *constructing a full Bayesian hypothetical prior* and using it to "reverse engineer" a value for $\mathfrak{c}(\mathbb{H},\mathbb{E})$ — as we did above, which led to the value $\mathfrak{c}(\mathbb{H},\mathbb{E}) = 1$. But, in this case, the reverse engineering leads to a value [$\mathfrak{c}(\mathbb{H},\mathbb{E}) = 1$], which is *outside* the range of values Douven assumes for $\mathfrak{c}(\mathbb{H},\mathbb{E})$.

In fact, there seems to be *no upper bound on* $\mathfrak{c}(\mathbb{H},\mathbb{E})$ that can be "reverse engineered" from Bayesian theory. I have a function for creating models satisfying the above constraints, but which require *arbitrarily large* values of $\mathfrak{c}(\mathbb{H},\mathbb{E})$.

In[ ]:=
```
model[c_] := PrSAT[{
    Pr[ℍ | 𝔼] > Pr[ℍ | ¬ 𝔼],
    Pr[ℍ | 𝔼 ∧ 𝔸] > Pr[ℍ | 𝔼 ∧ ¬ 𝔸],
    Pr[𝔸 | 𝔼] == Pr[𝔸 | ¬ 𝔼],
    Pr[ℍ | 𝔸] == Pr[ℍ | ¬ 𝔸],
    Pr[ℍ] == 1/2, Pr[𝔼] == 1/2, Pr[𝔸] == 1/2,
    Pr[ℍ | 𝔼 ∧ 𝔸] == (Pr[ℍ] × Pr[𝔼 | ℍ] + c) / (Pr[ℍ] × Pr[𝔼 | ℍ] + c + Pr[¬ ℍ] × Pr[𝔼 | ¬ ℍ])
  }, Probabilities → Regular, BypassSearch → True];
```

In[ ]:=
```
model[100]
```

Out[ ]=

$$\left\{\{\mathbb{A}\to\{a_2,a_5,a_6,a_8\},\ \mathbb{E}\to\{a_3,a_5,a_7,a_8\},\ \mathbb{H}\to\{a_4,a_6,a_7,a_8\},\ \Omega\to\{a_1,a_2,a_3,a_4,a_5,a_6,a_7,a_8\}\},\ \left\{a_1\to\frac{10828}{162409},\ a_2\to\frac{40528}{162409},\ a_3\to\frac{119097}{649636},\ a_4\to\frac{119097}{649636},\ a_5\to\frac{297}{649636},\ a_6\to\frac{297}{649636},\ a_7\to\frac{10828}{162409},\ a_8\to\frac{40528}{162409}\right\}\right\}$$