Epistemic Background

There are two main kinds of epistemic requirements. First, let’s examine these in the case of full (all-or-nothing) belief.

- **Correctness Requirements** (for full belief)
  - If \( p \) is false, then believing that \( p \) is incorrect.
  - Correctness requirements are *factive* — they involve relations between one’s attitudes and the *facts* [29, 23, 30].

- **Rational Requirements** (for full belief)
  - If one believes both \( p \) and \( \neg p \), then one’s beliefs are *structurally* irrational (viz., *incoherent*).
  - If one’s total evidence \( K \) counter-supports \( p \), then believing that \( p \) is *substantively* irrational.
  - Rational requirements are *non-factive* — they involve either (a) relations of (coherence) among one’s attitudes (*structural rationality*) or (b) relations between one’s total evidence \( K \) and one’s attitudes (*substantive rationality*) [31].

This distinction also applies to Bayesian epistemology...

Let us suppose that our agent is equipped with a *confidence ordering* \( \preceq \) over propositions. So that \( p \preceq q \) iff \( S \) is *at least as confident* in the truth of \( p \) as in the truth of \( q \).

And, \( p \succ q \) iff \( S \) is *strictly more confident* in \( p \) than in \( q \).

- **Correctness Requirements** (for comparative confidence)
  - If \( p \) is false and \( q \) is true, then \( p \succ q \) is incorrect.
  - Heuristically, correctness requirements involve *agreement with the attitudes of an omniscient agent* [16, 14, 20].

- **Rational Requirements** (for comparative confidence)
  - Intransitive \( \preceq \) relations are *structurally* irrational. E.g., the combination of attitudes \( \{ p \succ q, q \succ r, r \succ p \} \) is *incoherent*.
  - If one’s total evidence \( K \) supports \( p \) strictly more strongly than \( K \) supports \( q \), then \( q \succ p \) is *substantively* irrational.

Bayesians also like to talk about *numerical degrees of confidence* (credences). Indeed, this will be our main focus today...

Let us suppose that our agent has (at each time \( t \)) *degrees of confidence* (credences) represented by a function \( cr_t(\cdot) \).

The function \( cr_t(\cdot | E) \) will reflect the agent’s degrees of confidence (at \( t \)) — on the indicative supposition that \( E \) is true.

Finally, suppose our agent learns (exactly) \( E \) (with certainty) between \( t_0 \) and \( t_1 \). So, her transition from \( cr_{t_0}(\cdot) \) to \( cr_{t_1}(\cdot) \) reflects the upshot of learning (precisely) the content \( E \).

**Structural Bayesianism** involves a Trinity of Requirements [24].

1. **Synchronic (Non-Suppositional) Probabilism.** At each time \( t \), the agent’s (non-suppositional) credence function \( cr_t(\cdot) \) should obey the (Komlogorov) probability axioms.

2. **Synchronic (Suppositional) Ratio Formula.** At each time \( t \), the agent’s (suppositional) credence function \( cr_t(\cdot | E) \) should obey the ratio formula \( cr_t(\cdot | E) = \frac{cr_t(\cdot & E)}{cr_t(E)} \).

3. **Diachronic Conditionalization.** \( cr_{t_1}(\cdot) \) should equal \( \frac{cr_{t_0}(\cdot & E)}{cr_{t_0}(E)} \).
The idea behind Explanationism follows from (1)–(3).

4. **Learning & Supposing.** \( cr_t(\cdot) \) should equal \( cr_0(\cdot \mid E) \).

de Finetti [3] gave Dutch Book Arguments for both (1) and (2). Accuracy-dominance arguments for (1) are now popular [20]. One can also give an accuracy-dominance argument for (2) [12]. Lewis and others have aimed to adapt de Finetti’s argument for (2) into a “diachronic Dutch Book” argument for (3) [21].

Like de Finetti, I view (1) and (2) as the fundamental Bayesian principles governing the (structural) rationality of credences. I have always been more skeptical about the existence and nature of “diachronic coherence requirements.” Indeed, there seem to be many reasons to worry about (3) & (4) [13, 32, 25].

I will focus, primarily, on the synchronic requirements (1) & (2), and how they (allegedly) interact with Explanationism.

As Douven [5] explains, there have been various historical views regarding the proper formulation of Explanationism.

Rather than rehearsing Douven’s list of historical explications in detail, I will remain at a higher level of abstraction.

The idea behind Explanationism is that some “epistemic credit” should accrue to a hypothesis \( H \) in virtue of its being the best (or only) explanation of \( E \) (among the available alternatives \( \{H_j\} \)).

The question in which I am interested is: What is the best way to accommodate this basic Explanationist idea — of “credit” accruing to \( E \)'s best explanation — within a Bayesian framework?

van Fraassen [27] and Douven [5] maintain that a Bayesian should incorporate Explanationism by revising some of the basic requirements of Structural Bayesianism: (1)–(3).

I will focus on Douven’s proposal, since it is more precise, and it can be couched in purely synchronic terms [as a revision of (2)].

There are also substantive Bayesian requirements. In general, these are of the following generic form: If one’s total evidence at time \( t \) (\( K_t \)) is such and so, then \( cr_t(\cdot) \) should be thus and such.

Let \( \{H_1, \ldots, H_n\} \) be some partition of alternative hypotheses (putative explanations of \( E \) entertained by an agent (at time \( t \)).

**Substantive Bayesianism** (some example requirements)

- **The Principle of Indifference** [8]. If \( K_t \) does not favor any \( H_i \) over any \( H_j \), then \( cr_t(H_i) \) should equal \( cr_t(H_j), \forall i, j \).

- **The Principal Principle** [19]. If \( K_t \) entails that the objective chance of \( H_i \) is \( c \) (and \( K_t \) doesn’t contain/imply any inadmissible evidence), then \( cr_t(H_i) \) should equal \( c \).

- **The Requirement of Total Evidence** [17, 1, 30]. \( cr_t(H_i) \) should be equal to the evidential probability \( Pr(H_i \mid K_t) \).


Douven recommends that Bayesians revise the Ratio Formula (2) in such a way that the following alternative to Bayes’s Theorem is adopted (for hypotheses \( H_i \in \{H_1, \ldots, H_n\} \) and evidence \( E \)).

\[
\text{EXPL. } cr_t(H_i \mid E) = \frac{cr_t(H_i) \cdot cr_t(E \mid H_i) + c(H_i, E)}{\sum_{k=1}^{n} [cr_t(H_k) \cdot cr_t(E \mid H_k) + c(H_k, E)]}
\]

where \( c(H, E) \in [0, 1] \) is \( H \)'s “\( E \)-abductive credit score.”

And, \( c(H, E) > 0 \) iff \( H \) best explains \( E \) (o.w. \( c(H, E) = 0 \).

Note that if there is no best explanation of \( E \) among the \( \{H_k\} \), then EXPL reduces to Bayes’s Theorem (since all of the credit scores \( c(H_k, E) \) will be equal to zero in such a case).

Because EXPL leads to violations of Structural Bayesianism [viz., (2)], Douven discusses various ways a defender of EXPL might respond to de Finetti’s [3] Dutch Book argument for (2).

---

\(^1\)This definition can’t be quite right, since sometimes \( c(H, E) \) will need to be \( \geq 1 \) in order to emulate some Bayesian abductive updates. See Extras.
There is a more elegant way to capture the idea of “abductive credit,” which avoids the probabilistic incoherencies of [5, 27].

Note that there is something odd about thinking of EXPL as a structural requirement in the first place. EXPL has this form:

\[ \text{EXPL. } cr_t(H_i | E) \] should receive a boost — over and above the value prescribed by Bayes’s Theorem — just in case \((A_i) H_i \) is the best explanation of \(E \) (among the \( \{H_k\} \)).

Read literally, then, EXPL relates \( cr_t(H_i | E) \) to the fact that \( H_i \) best explains \( E \), which makes EXPL a correctness requirement.

In order for EXPL to be a structural requirement, it would have to be stated as a relation among the agent’s credences. To wit:

\[ \text{EXPL}_1. \quad cr_t(H_i | E) \] should receive a boost — over & above its Bayes’s Thm value — iff the agent is certain at \( t \) that \( A_i. \)

We also have to be careful to allow this “boost” to occur only once — presumably, when \( A_i \) is first learned. Otherwise, \( H_i \) will be “over-boosted.”

I’m not the first one to propose something like ERA. Climenhaga, Hartmann et. al., Lange, and Weisberg have all argued convincingly for similar principles [2, 18, 28, 4, 15, 9].

Roche & Sober [22] agree that ERA is the right formulation of Explanationism; but, they argue that ERA is false — viz., that \( A_j \) is never relevant to \( H_j \) (on the supposition of the evidence \( E \)).

I think Climenhaga [2] and Lange [18] do a pretty good job of responding to Roche & Sober’s skeptical argument [22].

The main thing I would add to Climenhaga’s and Lange’s trenchant responses Roche & Sober is the following point.

In order to refute R&S’s skepticism, all that is required is a single example in which \( A_j \) is relevant to \( H_j \) (given \( E \)).

I will close by discussing just such an example (involving Newton, Einstein, and the motion of Mercury), which is a well-known instance of The Problem of Old Evidence [6].
How would EXPL₁ handle this Old Evidence Problem?

Let us suppose that our agent became certain of \( A_2 \) in \( t_2 = 1915 \).

Then, according to EXPL₁, \( H_2 \) receives a “positive credit score” \( \mathcal{c}(H_2, E) > 0 \) and a boost in credence (at \( t_2 \))—over and above what Bayes’s Thm would recommend. Hence, EXPL₁ yields

\[
\mathcal{c}(t_2)(H_2 | E) > \mathcal{c}(t_2)(H_2).
\]

In this sense, according to EXPL₁, \( E \) confirms \( H_2 \) in 1915.

I see (at least) two problems with this application of EXPL₁.

- It is misleading to say that \( E \) is what confirms \( H_2 \) in 1915. It is \( A_2 \) — by ensuring \( \mathcal{c}(H_2, E) > 0 \) — that is doing the work.

- In addition to the machinery of probability theory [modulo (2)], Douven also needs a “theory of credit scores.” In this sense, the (Bayesian) ERA approach is more parsimonious. [See Extras for a detailed formal example, which brings this out.]

The simplest formal example of Douven v. Bayes involves a 2-element partition \( \{ H_1 = H, H_2 = \neg H \} \), evidence \( E \), and an abductive claim \( A \) asserting that \( H \) explains \( E \) better than \( \neg H \).

Our Bayesian agent will begin with a prior \( \mathcal{c}(t_0)(\cdot) \) over the algebra generated by the three atoms \( H, E, A \). Between \( t_0 \) and \( t_1 \), they will learn \( E \), and between \( t_1 \) and \( t_2 \) they will learn \( A \).

I will assume \( \mathcal{c}(t_0)(\cdot) \) satisfies the following six (6) constraints.

1. \( \mathcal{c}(t_0)(\cdot) \) is regular.
2. \( E \) (initially) confirms \( H \) \( \mathcal{c}(t_0)(H | E) > \mathcal{c}(t_0)(H | \neg E) \).
3. \( \text{ERA} \) applies \( \mathcal{c}(t_0)(H | E & A) > \mathcal{c}(t_0)(H | E \& \neg A) \).
4. \( E \) is (a priori) irrelevant to \( A \) \( \mathcal{c}(t_0)(A | E) = \mathcal{c}(t_0)(A | \neg E) \).
5. \( A \) is (a priori) irrelevant to \( H \) \( \mathcal{c}(t_0)(H | A) = \mathcal{c}(t_0)(H | \neg A) \).
6. \( \mathcal{c}(t_0)(H) = \mathcal{c}(t_0)(E) = \mathcal{c}(t_0)(A) = \frac{1}{2} \).

There are many priors that satisfy (i)–(vi). I will choose a relatively simple one for illustrative purposes.