Chalmers’s Frontloading Argument for A Priori Scrutability

RAM NETA

David Chalmers’s new book Constructing the World is a brilliant defence of a comprehensive, systematic philosophy. To bite off a part of it for rumination in these comments, I’ll begin with some historical context.
Nearly two decades ago, Chalmers gave an influential argument that went roughly like this:

1. Zombies are conceivable.
2. Therefore, the totality of microphysical facts does not \textit{a priori} entail the facts about consciousness (from 1).
3. Therefore, there is no metaphysical reduction of consciousness to the microphysical facts (from 2).

Each step of that argument has been challenged. But one of the most profound challenges to date was that offered by Ned Block and Robert Stalnaker in 1999: they argued against the second step of the argument (from 2 to 3). Their argument went like this:

4. The totality of microphysical facts does not \textit{a priori} entail ordinary facts about water boiling.
5. Nonetheless, ordinary facts about water boiling are metaphysically reducible to the microphysical facts.
6. Therefore, \textit{a priori} entailment is not necessary for metaphysical reduction (from 4 and 5).

In their famous 2001 reply to Block and Stalnaker, Chalmers and Frank Jackson argued against (4). Or rather, they argued against something stronger than (4) itself, since (4) is both obviously true, and also not as strong as what Block and Stalnaker need to claim to challenge Chalmers’s original argument effectively. Obviously true, because of course there are facts about water boiling that are not \textit{a priori} entailed by the totality of the microphysical facts. For instance, the fact that no boiling water consists of immaterial stuff is not \textit{a priori} entailed by the totality of the microphysical facts, because the totality of the microphysical facts does not rule out the existence of lots of immaterial types of water. The fact that there is water boiling \textit{right here} is another fact that is not \textit{a priori} entailed by the totality of the microphysical facts, since the totality of those facts does not entail anything indexical. And – at least according to those who take there to be an explanatory gap between the physical and the qualitative – the fact that cold water is more pleasant to drink than boiling water is also not \textit{a priori} entailed by the totality of the microphysical facts, since the totality of those facts does not entail anything about that qualitative fact. So (4) is easy to verify. But (4) is not as strong as what Block and Stalnaker need to make their case against Chalmers. To make their case against Chalmers, Block and Stalnaker need to defend the stronger claim.

7. The totality of microphysical, indexical, qualitative, and ‘that’s all’ facts does not \textit{a priori} entail the ordinary facts about water boiling.

If (7) is true, then Block and Stalnaker could appeal to the conjunction of (7) and (5) to argue for (6), and thereby undermine Chalmers’s original
argument from (2) to (3). So Block and Stalnaker’s challenge to Chalmers is well served if (7) is true.
In fact, even if (7) is not true, Block and Stalnaker could fall back on

(8) Some compact totality of facts (including the microphysical, indexical, qualitative, and ‘that’s all’ facts) does not \textit{a priori} entail the ordinary facts about water boiling, even though the ordinary facts about water boiling are metaphysically reducible to that compact totality of facts.

A ‘compact’ totality of facts is one that, in Chalmers’s words, ‘involves only a limited class of concepts and that avoids trivializing mechanisms such as coding the entire state of the world into a single number’ (Chalmers 2012, xiv). If (8) is true, then metaphysical reduction does not require \textit{a priori} entailment, and that is just what Block and Stalnaker set out to show. So, if Chalmers’s argument from (1) to (2) to (3) is to avoid any form of the Block/Stalnaker objection, Chalmers must argue that (8) is not true. And that is what he does in his new book.

Actually, what Chalmers does is not simply try to cast doubt on (8). Rather, Chalmers offers a sustained and ingenious positive argument in support of

(9) The totality of facts is \textit{a priori} entailed by some compact totality of facts.

This is the thesis that Chalmers calls ‘\textit{A Priori Scrutability’}. In a book full of fascinating arguments for provocative claims, \textit{A Priori Scrutability} is the thesis at the foundation of the whole project. Chalmers’s chief argument for that thesis is what he calls ‘the argument from frontloading’. The argument has two versions, as follows:

Suppose that one has conditional knowledge that if PQTI, then M. Suppose also that this knowledge is justified by some empirical evidence E. Then one is plausibly in a position to know that if PQTI&E, then M. Furthermore, E will not play an essential role in justifying this conditional knowledge: there is no need for it to do so, as E is built into the antecedent, and its justifying role in reaching the conditional conclusion that M from the supposition of PQTI can be played just as well by supposing it as by believing it. Perhaps, the knowledge that if PQTI&E, then M is itself justified by some further evidence, but then one can repeat the process by conjoining this evidence to the antecedent. If one repeats this process for all relevant empirical evidence, one will eventually end up with a large conjunction F of evidence statements such that one can know that if PQTI&F, then M without justification from any empirical evidence. That is, one can know ‘If PQTI&F, then M’ \textit{a priori}.

This reasoning is especially natural in a Bayesian framework. Suppose that ‘cr*(M|PQTI) is high, and that this credence is justified by some class
of empirical evidence sentences E. Then \( \text{cr}^*(\text{M|PQTI&E}) \) will also be high. . . . if acquiring total evidence E enables one to have a high rational credence \( \text{cr}^*(\text{M|PQTI}) \), then even before acquiring evidence E, one is in a position to have a high rational credence \( \text{cr}^*(\text{M|PQTI&E}) \). So it is plausible that E plays no essential role in justifying one’s high rational credence \( \text{cr}^*(\text{M|PQTI&E}) \). By repeating this process, one will end up with a large class of evidence sentences F such that a high rational credence \( \text{cr}^*(\text{M|PQTI&F}) \) is justified \emph{a priori}.’ (Constructing the World, 161)

I don’t know whether the conclusion of the frontloading argument is true. But whether or not it is true, I believe that the frontloading argument, in both of the two versions that Chalmers offers, is not compelling. In the remainder of these comments, I will explain why.

I will begin with the Bayesian version of the argument. This version of the argument depends upon the following principle, according to which an epistemic agent is less than fully rational if all of the following three things are true:

(i) at time \( t_1 \), her conditional credence \( (H|E) = n \)
(ii) at time \( t_1 \), her total evidence changes simply by the addition of E
(iii) immediately after time \( t_1 \), her unconditional credence \( (H) \neq n \).

Let’s call this ‘the principle of conditionalization’. Now notice that the principle of conditionalization, as I just stated it, does \emph{not} enjoin an agent of whom (i) and (ii) are true to avoid letting (iii) be true as well: that is, the principle of conditionalization does not enjoin an agent to perform the action generally known as ‘conditionalizing’. The principle does \emph{not} say: ‘if your conditional credence \( (H|E) = n \), and your total evidence changes simply by the addition of E, then see to it that your unconditional credence \( (H) = n \).’ An agent of whom (i) and (ii) are both true could know, for instance, that she was being irrational in letting her conditional credence \( (H|E) = n \). In such a case, it is not clear that she should, upon acquiring evidence E, let her unconditional credence \( (H) \neq n \). In such a case, she is at least rationally permitted – and perhaps even rationally required – to let her unconditional credence \( (H) \neq n \). This is not a counter-example to the principle of conditionalization as I stated it above, for it is a case in which the epistemic agent is less than fully rational – though in this case her irrationality consists not in meeting condition (iii) when she meets conditions (i) and (ii), but rather in her specific way of meeting condition (i). This is one circumstance in which an agent of whom (i) and (ii) are both true could rationally allow (iii) to be true as well: her irrationality consists not in her allowing (iii) to be true of her, but rather in her having (i) be true of her. In other words, the principle of conditionalization does not require us always to conditionalize; rather, it simply says that our failure to conditionalize is a sufficient condition of our being less than fully rational – even if it is sufficient in a particular case only
because of some antecedent irrationality the mitigation of which requires us, in this case, not to conditionalize.

Such cases arise frequently. One way they have arisen in the history of science is when we discover that we had assigned an extremal prior probability to some hypothesis that is empirically contestable, e.g. Euclidean propositions about spatial relations, or pre-Relativistic propositions about temporal relations. In such cases, rationally revising our beliefs requires us to let our unconditional credence (H) when we acquire evidence E be unequal to our conditional credence (H|E) just before acquiring E, and that is because we were irrational in having an extremal prior probability for the relevant H.

Bayesians typically accept an a priori argument of the following form for the claim that our prior probabilities in non-logical truths should all be non-extremal:

\begin{align*}
(10) & \text{If we assign an extremal prior probability to } H, \text{ then there is no evidence that we can acquire such that conditionalization on that evidence would alter our rational credence in } H. \\
(11) & \text{There is no empirical } H \text{ such that evidence cannot change our rational credence in } H. \\
(12) & \text{The only way for evidence to change our rational credence in } H \text{ is by conditionalization on that evidence.} \\
(13) & \text{Therefore, for any empirical } H, \text{ it is less than fully rational to assign an extremal prior probability to } H. \text{ (from 10 to 12)}
\end{align*}

This argument is valid, and (10) is certainly true. I suspect that (11) is true also, but I'm not sure of that, since I've never heard a good non-inductive argument for (11). Although I do not know of any argument for (12), the dynamic Dutch Book argument that Lewis gives for the principle of conditionalization is an argument to the effect that, whenever evidence changes our rational credence in H, it must do so in a way that gives the same result as conditionalization. And that is really all that the Bayesian needs for the argument. In any case, the argument above is very likely sound, and I accept it.

So let's grant the Bayesians that this argument imposes an a priori constraint of rationality on our priors. Not that this argument could be appreciated in advance of receiving any evidence: a creature that has never received any evidence might fail to have any concepts at all, and so not be in a position to understand, let alone feel the force of, this a priori argument against extremal priors. But this a priori constraint on our priors need not be understood or appreciated by such a creature. There are, of course, other a priori constraints on our priors that might not be understood or appreciated in advance of receiving any evidence: for instance, the constraint of synchronic probabilistic coherence is one such. The a priori argument for this constraint is not one that is generally (or perhaps even possibly) understood or appreciated by creatures who have never received any evidence – but this does not rob it of its normative force. So the Bayesian will, I assume, grant that there
are rational constraints on our prior probabilities, and that at least some of these rational constraints cannot be appreciated or understood by agents who have not yet received any evidence.

This need not require the Bayesian to make any concessions to epistemic externalism: the Bayesian could instead say that the principle of conditionalization imposes a rational requirement on the evolution of our credences in light of new evidence only once we (who are subject to that requirement) have already gained some substantial a priori knowledge – an achievement that requires us to have received a great deal of evidence already. This is not an implausible view: the dynamic Dutch Book argument for the principle of conditionalization shows that an agent who violates that principle is thereby committed to the fairness of a set of bets that jointly guarantee a loss; but perhaps commitment to the fairness of a bet can only be undertaken by a creature who already has some substantial a priori knowledge. Only against the background of such knowledge can a creature do anything that would count as issuing such a commitment. If that is the case, the dynamic Dutch Book argument might plausibly be understood to show only that agents who have the requisite body of a priori knowledge are rationally required not to violate the principle of conditionalization.

Might there be other constraints of rationality on our priors besides those a priori constraints I have already mentioned – perhaps even a posteriori constraints on our priors? The idea of an a posteriori constraint on our ultimate priors might seem to be contradictory, but I will now argue that it is not.

Consider two epistemic agents – Harry and Sally – who are both perfect conditionalizers, and who are subject to precisely the same evidential history (modulo indexicals), but who start from different ultimate priors. For the sake of simplicity, let’s assume that their evidential history consists exclusively of 12 coin tosses. Harry’s ultimate priors assign equal probability to each of the $2^{12}$ possible outcomes of the coin tosses, whereas Sally’s ultimate priors do not. Now let’s suppose that they each witness the following 11 tosses:

```
HHHHHHHHHHH
```

Neither of them has yet witnessed the 12th toss. At this point, what credence will Harry assign to the 12th toss being H and what credence will he assign to its being T? Recall that Harry is a perfect conditionalizer, so he will conditionize on the evidence that he’s received up to this point in determining his credence for each of H and T for the 12th toss. Here’s how that will go for Harry:

\[
\begin{align*}
P(H_{12}) &= P(H_{12} | HHHHHHHHHHH) = \\
P(H_{12} &\& HHHHHHHHHHH) | P(HHHHHHHHHHH) = 0.5
\end{align*}
\]

And similarly, $P(T_{12}) = 0.5$
Thus, Harry will assign equal credence to the 12th toss being H and to its being T: all that the first 11 tosses do for his credence function is rule out the possibility of any distribution in the first 11 tosses other than HHHHHHHHHHHH. Those first 11 tosses do not count, for Harry, as evidence that the 12th toss will be H. Sally, in contrast, can treat the results of the first 11 tosses as very strong evidence that the 12th toss will also be H, since her ultimate priors do not assign equal probability to each possible distribution of the 12 coin tosses. Clearly, Sally’s ultimate priors allow for more efficient learning from evidence than Harry’s ultimate priors do.

There are well known a priori arguments for the conclusion that any set of non-extremal priors will allow for evidential learning in the long run. Unfortunately, as a matter of empirically known fact, no human being ever gets to live into the long run: our mortality limits the duration of our evidential history. And while Harry’s ultimate priors might allow Harry to enjoy evidential learning in the long run, they will not allow Harry to enjoy evidential learning if his evidential history is confined to 12 coin tosses. And what holds true of Harry’s ultimate priors also holds true of the ultimate priors of any mortal epistemic agent: if their ultimate priors do not allow that agent to enjoy evidential learning over the course of her limited evidential history, then those ultimate priors are just as irrational as extremal valued priors that also do not allow her to enjoy evidential learning. In other words, it is a matter of empirically known fact that some non-extremal ultimate priors do not allow for evidential learning over the course of a human lifetime. If extremal priors are irrational because they do not allow for evidential learning in the long run, then, I claim, even these non-extremal priors that do not allow for evidential learning over the course of a human lifetime are also irrational. But whereas it is a priori that extremal priors do not allow for evidential learning ever, it is a posteriori that some non-extremal priors do not allow for evidential learning over the course of a human lifetime. And so this is an a posteriori constraint on the ultimate priors of mortal epistemic agents. Of course, this is not a constraint that we can appreciate or understand prior to receiving any evidence – but then, by the Bayesian’s own admission, neither is the a priori constraint one that we can appreciate or understand prior to receiving any evidence.

Again, this need not require the Bayesian to make any concessions to epistemic externalism: the Bayesian could instead say that the principle of conditionalization imposes a rational requirement on the evolution of our credences in light of new evidence only once we have already gained some substantial a posteriori knowledge. And again, this is not an implausible view: the dynamic Dutch Book argument for the principle of conditionalization shows that an agent who violates that principle is thereby committed to the fairness of a set of bets that jointly guarantee a loss: but perhaps commitment to the fairness of a bet can only be undertaken by a creature who already has some substantial a posteriori knowledge. Only against the
background of such knowledge can a creature do anything that would count as issuing such a commitment. If that is the case, the dynamic Dutch Book argument would only show that agents who have the requisite body of a posteriori knowledge are rationally required not to violate the principle of conditionalization.

I conclude, then, that there are a posteriori constraints of rationality on our ultimate priors – constraints determined by whether these ultimate priors allow for evidential learning over the course of our lifetime. Perhaps there are also a posteriori constraints of rationality that are determined by the degree to which those priors allow us to enjoy efficient learning, given the normal range of evidence to which we have access. But, whether or not there are such additional a posteriori constraints, I believe I have shown that there are some a posteriori constraints of rationality on our ultimate priors. And so even if Chalmers is right to say that our rational conditional credence in T (the totality of facts), given the conjunction of C (the compact scrutability base of facts) and our total evidence, must be high, and that it must be high independently of the total evidence upon which we conditionalize, it does not follow from this that our rational credence (T|C&our total evidence) must be high a priori. Conditionalization is not the only a posteriori rational constraint on our credence function.

I have now said why I don’t find the Bayesian version of the frontloading argument compelling. Let me now turn to the first version of the argument. Recall that the crucial move in that argument is this: if the inference from C to T is supported by E, then the inference from the conjunction C&E to T will not itself be supported by E. Why should we accept this conditional? Chalmers devotes a couple of paragraphs to answering this question:

These frontloading principles have strong intuitive support. One can argue for the simple frontloading principle as follows. Given that E justifies M, then one could in principle (i) suspend judgment concerning E, (ii) suppose (for the purposes of conditional reasoning) that E, (iii) conclude (under this supposition) that M, with justification provided by E’s support for M, and (iv) discharge the supposition, yielding a justified conditional belief in M given E. This conditional belief is justified even though one has suspended judgment concerning E, so that E played no non-suppositional role in its support. So the conditional belief in M given E is justified independently of E.

The main worry about Chalmers’ reasoning in the preceding paragraph concerns step (iii), and Chalmers anticipates this worry in the very next paragraph:

The main question concerns step (iii): could it be that E’s support for M itself somehow depends on E, in a way such that suspending judgment about E also undermines the epistemic connection between E and M?
That would be at least odd. Typically, if P’s support for Q itself depends on support from some further claim R, then one can combine these elements of support, yielding a combined support by P&R for Q that does not depend on R’s support in this way. On the face of it, in this fashion one could combine all the ways that E provides support into a single support relation that does not depend on E.

Now, I agree that what Chalmers here says is ‘typically’ the case is indeed typically the case. But what Chalmers needs to claim here, for the sake of his argument, is that it is invariably true. And now I’d like to give some grounds for doubt that it is invariably true, by producing what seems to me to be a possible counter-example.

Consider the hypothesis that you are not at this moment dreaming. If you are justified in believing that hypothesis, then you are justified in believing it empirically. That is not a hypothesis that you could be a priori justified in believing, since your justification for believing must involve, at least to some extent, the experiences that you are having now: if those experiences were oddly dreamlike, rather than routinely philosophy paperlike, then you would not have as much empirical justification for believing the hypothesis as you do now. But there is no single part of your current experience that serves as a crucial piece of evidence that you are not dreaming. Of course, you can pinch yourself in an effort to prove to yourself that you are not dreaming, but the pinching sensation alone does not justify for you the hypothesis that you are not dreaming; it is rather the total body of experience of which the pinching sensation is a part that justifies that hypothesis for you. Your justification for believing that you are not dreaming has to do with how the total body of your empirical evidence all fits together. As Descartes puts it at the end of the Sixth Meditation, in discussing the difference between dreams and waking consciousness:

I now find a very marked difference between the two states, in respect that our memory can never connect our dreams with each other and with the course of life, in the way it is in the habit of doing with events that occur when we are awake. . . . But when I perceive objects with regard to which I can distinctly determine both the place whence they come, and that in which they are, and the time at which they appear to me, and when, without interruption, I can connect the perception I have of them with the whole of the other parts of my life, I am perfectly sure that what I thus perceive occurs while I am awake and not during sleep.

So you are justified in believing the hypothesis

W: I am not dreaming now

on the basis of your total empirical evidence E. But what justifies you in believing the conditional ‘If E, then W’? If justification is closed under
known logical consequence, then you must be somehow justified in believing this conditional so long as you are justified in believing W, since the conditional obviously follows from W. But what justifies you in believing it? Are you justified in believing it \textit{a priori}? It is hard to see how you could gain \textit{a priori} access to a conditional that concerns the body of empirical evidence that you have \textit{right now}, and the hypothesis that you are not dreaming \textit{right now}. Perhaps you have \textit{a priori} justification for believing a conditional of the form:

\[(G) \text{ If, at a particular time } t, \text{ I am having a more-or-less normal body of empirical evidence, then I am not dreaming at } t.\]

But even if you do have \textit{a priori} justification for believing G, that still does not suffice for you to have \textit{a priori} access to the conditional ‘If E, then W’, since it is not \textit{a priori} that your \textit{current} body of empirical evidence E is a \textit{normal} body of empirical evidence, in whatever sense of ‘normal’ is at issue in G. If you are justified in believing that E is a normal body of empirical evidence, then E itself must be part of what justifies you in that belief, and so your justification for believing that E is a normal body of empirical evidence is itself an empirical justification.

So, even if you have \textit{a priori} justification for believing G, you still don’t have \textit{a priori} justification for believing ‘If E, then W’. And if you do not have \textit{a priori} justification for believing G, then on what basis can we claim that you have \textit{a priori} justification for believing ‘If E, then W’?

Of course Chalmers may suggest such a basis: namely, the frontloading argument! But what is at issue now is whether step (iii) of that argument is invariably true: we cannot appeal to the frontloading argument to justify one of its own steps. Perhaps Chalmers would reply by claiming that this single dreaming example is too isolated and controversial for it to bear much weight against the seemingly intuitive step (iii) of the frontloading argument. So perhaps it is best to attack the intuitive support of that step directly. What lends (iii) its intuitive support?

I believe that what lends (iii) its intuitive support is simply that (iii) is true for all cases of deductive support, which are the cases that are most salient to many philosophers. If E entails M, then even if we merely suppose E and deduce M from E within the scope of that supposition, we can then detach the supposition and conclude that the conditional ‘If E, then M’ is true \textit{a priori.} But why think that this works for non-deductive support? Suppose that E evidentially supports M. Must it then be true that even if we merely suppose E we can then infer M from E within the scope of the supposition? Not if our ability to infer M from E itself depends upon our having justification for believing that E is true. For Chalmers to defend step (iii) of his frontloading argument, he would have to show that our ability to infer a conclusion from a body of evidence does not depend upon our having justification for believing that body of evidence is true. And I do
not see that this can be done. In fact, I suspect that one lesson of Goodman’s ‘grue’ puzzle is that the support that a hypothesis gets from our total evidence depends upon our justification for believing that evidence. An insentient rational being who merely contemplated the ‘green’ induction and the ‘grue’ induction would not be able to see why one induction is better than the other given our total body of evidence – but we, who aren’t merely supposing that evidence to be true, but are justified in believing it to be true, can plainly see that one is better supported by that evidence than the other, even if there is no formal account of inductive support to explain why this is.

I conclude that you do not have a priori justification for believing ‘If E, then W’. But, given that you are justified in believing W, and that justification is closed under known logical consequence, it follows that you do have justification for believing ‘If E, then W’. Since you don’t have a priori justification for believing ‘If E, then W’, it follows that you have a posteriori justification for believing ‘If E, then W’. If (as is typical) ‘If E, then W’ is not justified by some single bit of empirical evidence that you have, then it is justified by your total evidence E. And so, I conclude, you are typically justified in believing the conditional ‘If E, then W’ on the basis of your total evidence E.

This is my counter-example – or at least possible counter-example – to the general principle on which the non-Bayesian version of Chalmers’s frontloading argument relies, viz., that the support that evidence lends a hypothesis cannot itself depend on that very same evidence. If this general principle is false, then the non-Bayesian argument does not work. I’m not sure that the general principle is false, but I’ve just attempted to cast doubt on it by appeal to a possible counter-example.

The central thesis of Chalmers’s magnificent book is A Priori Scrutability, viz., that the totality of facts is a priori entailed by some compact totality of facts. And his central argument for that thesis is the frontloading argument. I have attempted to show that neither version of that argument is compelling. If A Priori Scrutability is true, then that still remains to be shown.1

University of North Carolina
Chapel Hill, NC 27599
neta@email.unc.edu

1 Thanks to Matt Kotzen, John Roberts, Robert Smithson, Michael Titelbaum, and especially David Chalmers, for discussion.