

REASONING ABOUT REASONING*

TYLER BURGE

Mutual expectations cast reasoning into an interesting mould. When you and I reflect on evidence we believe to be shared, we may come to reason about each other's expectations. I develop expectations about your expectations, expecting them to be expectations about mine. The double-mirror aspect of such situations suggests that there may in some sense be no stopping. Of course, we do run out of expectations. But to some this seems to result merely from inability to understand sufficiently complicated nestings, rather than from lack of *reason* to acquire each of infinitely many nested beliefs.

Limitless nesting of psychological "contents" has played an important role in accounts of communication, convention and related topics. But there have been few attempts to analyze precisely how the infinity of contents might be generated. An exception is David Lewis' account in *Convention* pp. 52-57. This account contains significant errors which are easily overlooked because of the complexity and subtlety of the subject matter. Fortunately, they are also easily corrected, at least within the general framework of the account. The purpose of this note is to identify and correct the oversights.

A key to the discussion is the term 'have reason to believe.' The term is used by Lewis in such a way that if a person has reason to believe something and does not already believe it, he would be extremely irrational to disbelieve it and he should come to believe it with only a little reflection (pp. 53, 59, 63). Having reason is understood to vary with one's inductive standards and background information.

The reasons normally “had” in the situations we shall be discussing are inductive reasons. They therefore admit of uncertainty. In what follows I shall sometimes speak of one’s reason to believe ϕ as making ϕ probable for one to a certain specified degree. It is unimportant to my purpose whether inductive reasoning be actually, or even ideally, explicable in terms of numerical probabilities. The appeal to probability is merely a way of making applications of the following point picturesque: When one has reason to be less than fully certain about each of a plural number of epistemically independent theses the assumption of each of which is needed to justify a conclusion, one’s certainty about the conclusion (relative to those assumptions) should be less than one’s certainty about any one of the assumptions. It is common to assume that if numerical probabilities are assigned, the probability of the conclusion (relative to the reasons for believing it) is no more than the product of the probabilities assigned to those reasons, assuming the reasons independent.¹

Lewis illustrates his theory of interpersonal reasoning by reference to a particular case. We imagine that *state of affairs A* is any sort of evidence for the conclusion that we will meet at a given spot tomorrow – for example, a signal of some kind. I shall call the state of affairs of our meeting at that spot and time *state of affairs B*. Lewis makes the following assumptions, and I shall adopt them:

- (1) You and I have reason to believe that some state of affairs A holds.
- (2) Each of us is such that if he had reason to believe that A held, then he would thereby have reason to believe that you and I have reason to believe that A held.
- (3) Each of us is such that if he had reason to believe that A held, then he would thereby have reason to believe that B would hold.
- (B) Each of us has reason to believe that if he would have reason to believe B held granted that he had reason to believe A held, then the other would have reason to believe B held granted that *he* had reason to believe A held.

Lewis claims that these four premises imply

- (4) Each of us is such that if he had reason to believe that A held, then he would thereby have reason to believe that each of us has reason to believe that B would hold.

REASONING ABOUT REASONING

He also claims that they imply

- (5) Each of us is such that if he had reason to believe that A held, then he would thereby have reason to believe that each of us would have reason to believe that the other has reason to believe that B would hold

and so on, with indefinitely many iterations in the consequent.²

These claims are mistaken. We can most easily bring out why by representing simplified versions of the relevant sentences in symbolic, abbreviated form:

- (1') $R(I,A)$
 (2') $R(I,A) \rightarrow R(I,[R(I,A)]) \ \& \ R(I,[R(Y,A)])$
 (3') $R(I,A) \rightarrow R(I,B)$
 (B') $R(I [R(I,A) \rightarrow R(I,B), \rightarrow R(Y,A) \rightarrow R(Y,B)])$
 (4') $R(I,A) \rightarrow R(I [R(I,B)]) \ \& \ R(I, [R(Y,B)])$

'R' stands for 'have reason to believe'; 'I,' for 'I'; 'Y,' for 'you.' The brackets are intended for ease of reading. The arrow stands for 'if ... then (thereby) ...' — except in its occurrence as main connective in the content clause of (B'), where it stands for ordinary 'if ... then ...': the difference in the meaning here can be slurred without harm. I assume that the analog of *modus ponens* holds for the arrow. We do not need to give a semantics for the notation. Whereas the argument from (1') through (4') represents only the first person point of view, (1)-(4) incorporate an analogous argument from the second person viewpoint. Our remarks will apply (by analogy) to both arguments.

Using (1') and (3') we can detach

- (a) $R(I,B)$.

To get the first conjunct in the consequent of (4'), we need the iteration principle

- (6') $R(I,C) \rightarrow R(I, [R(I,C)])$

where 'C' stands for any content clause. Let us grant (6') for the present. This together with (a) gives us

- (b) $R(I, [R(I,B)])$.

Using (1') and (2') we can detach and simplify to

- (c) $R(I, [R(Y,A)])$.

Now my reason described in (c) is inductive. Suppose it has a probability strength of 65%. My reason described in (B') is also

inductive. Suppose it has a probability strength of 66%.

Using (B') and (3'), one is tempted to try a sort of detachment that would yield

$$(d) \quad R(I, [R(Y,A) \rightarrow R(Y,B)]).$$

But such a detachment would be within the scope of 'R' – within the first person point of view. And so far, we have assumed only (3'), not that I have reason to believe (3').

Let us grant the following for the sake of the argument:

$$(7') \quad R(I,C) \ \& \ (R(I,C) \rightarrow R(I,D)) \rightarrow R(I, [R(I,C) \rightarrow R(I,D)])$$

where 'C' and 'D' stand for any content clauses. The idea is that if I have reason to believe some state of affairs obtains, and that state of affairs gives me reason to believe that another state of affairs obtains, then I have reason to believe that this evidential or reason-given relationship holds for me.

With (7'), (1') and (3') we can derive

$$(e) \quad R(I, [R(I,A) \rightarrow R(I,B)]).$$

The status of the reason described in (e) is a deep problem. Hume argued that such reasons were clearly not deductive and that if they were inductive, then induction could not be justified. I do not want to embroil myself in this matter. But I shall suppose, what seems plausible, that my reason described in (e) has a probability strength of less than 100%.

We are now in a position to diagnose the sources of Lewis' mistake. My reason in (B') has a probability strength of only 66%, and any premises I conjoin with that reason may carry their own degrees of uncertainty. Unless such premises are 100% probable (or epistemically derivative), the conjunction will have less strength than any of the conjuncts. Thus we may derive (d) from (B') and (e) *only if* the product of the probability strengths of my reasons – those reasons described in (B') and (e) is sufficiently high. If it is, we obtain (d). The same point applies in conjoining (d) and (c) to get

$$(f) \quad R(I, [R(Y,B)]).$$

(This derivation is more clearly problematic than that of (d) because its interpretation does not hinge on the Humean issues surrounding (e).) In fact, on our assumptions about probability strength in (c) and (B'), we *cannot* derive (f). For the reason described in (d) has a strength of less than 66%; that in (c) has a strength of 65%; the

REASONING ABOUT REASONING

reasons are (reasonably) independent epistemically; so their product is less than 50%. In ordinary circumstances, this situation might be regarded as one in which I lack the reason described in (f). If my inductive support were strong enough to justify (f), of course, we could conjoin (f) with (b), and thereby have (4').

The moral is that the truth of (4) depends on the combined *strength* of our reasons regarding three matters: the other's appreciation of the evidence ((2')), the other's inductive standards and background information ((B')), and our own inductive reasons ((7') and actually (6') as well).

How will iteration affect matters? (1), (2), (4) and (B) (or analogs, cf. note 2) do not entail (5) for the same reason that the previous alleged entailment failed. But if (B'), (7') and (c) are strong enough to produce (4'), then by relying on (4') (and an analog of (B') which is of at least equal probability strength), one can derive the appropriate representation of (5) – on one condition. The reasons generated by (6') and (7') must not become weaker as these principles are reapplied. This condition seems no less plausible than the iteration principles themselves.

Thus Lewis' entailment claims can be corrected by making explicit the principles (6') and (7'), and by requiring that the combined reasons cited in the premises be sufficiently strong. Of course, even this account might be doubted. Not all epistemologies would accept (6') and (7'). It might be thought that one could have first-order reasons but lack the information or subtlety to have higher-order reasons. Moreover, there is the bothersome Hume, who had inductive reasons to believe the sorts of things everyone else believes, but who doubted that he had those reasons. On Lewis' account of 'have reason to believe' (accepting (6') and (7')), we would have to conclude that Hume was "extremely irrational." So doing would be a mite severe. Whether this difficulty is best accommodated by minor refinements in the interpretation of 'have reason to believe' or by important changes in the iteration principles, I leave as an open question.

UNIVERSITY OF CALIFORNIA, LOS ANGELES
LOS ANGELES, CALIFORNIA 90024
USA

TYLER BURGE

NOTES

- * I am indebted to Gregory Kavka for remarks on an earlier version.
- 1 The assumption of the independence of justifications must be seen as an idealization. It has been argued that it is an unfruitful and fundamentally misleading idealization. Since our appeal to probability strength is purely heuristic, this issue can be set aside.
- 2 (1)-(5) are formed from like-numbered sentences in *Convention* (Harvard University Press, Cambridge, 1969), pp. 52-53, by replacing 'indicates' in the latter sentences by Lewis' definition of the term. A *indicates* to someone x that --- if and only if, if x had reason to believe that A held, x would thereby have reason to believe that ---. (B) is the relevant instance of Lewis' assumption that you and I have reason to believe that we share the same background information and inductive standards, at least nearly enough so that A will indicate the same things to both of us. Actually Lewis claims that instances of this assumption would be needed to get (5) and more complicated iterations. Thus it is (1)-(3) and relevant *analogues of (B)* which are alleged to imply (5). Ignoring this point will not be crucial to our criticism.