of his alternatives. (2) It may be that Z and U are two of the alternatives for a certain subject. Then the four-place relation serves to cross-identify things between the worlds of different alternatives.

The worlds of different alternatives need not be different worlds, nor need they be different from the subject's own world. It may happen, then, that a thing is or isn't cross-identified by acquaintance with itself. Consider, for instance, the situation in world 2 of our example: X for U is not a counterpart by acquaintance of X for V.

A subject of attitudes can scarcely fail to be intimately acquainted with himself, and the subject's alternatives will be likewise self-acquainted. So we will have cases that are reflexive in the first two places and in the last two. Normally, perhaps invariably, when U is one of Z's alternatives, then U for U is a counterpart by acquaintance of Z for Z. It is an interesting question whether we can have reflexivity in the last two places only: T for U is a counterpart by acquaintance of Z for Z, where T and U are not identical. Such would be a case where some of Z's major channels of self-acquaintance are relations which he could have born to something other than himself: he watches himself in the mirror he mistakes for a window, he follows his own trail, or what have you. Shall we permit such cases? Or shall we rule them out, giving identity predominant weight among channels of self-acquaintance?[23] I know of no evidence that we have settled this question. Nor need we settle it; we can allow that different versions of the four-place counterpart-by-acquaintance relation go different ways.

23 This question is relevant to Hintikka's proposal, in his "On Attributions of 'Self-Knowledge,'" to represent self-knowledge using formulas that are to be interpreted using cross-identification by acquaintance. The proposal succeeds only if we do give identity predominant weight. At this point I am indebted to discussions with Mark Johnston.

23

Why conditionalize?

INTRODUCTION (1997)

This paper presents what is nowadays called the 'diachronic Dutch book argument'. I wrote it in 1972 as a handout for a course, with no thought of publication. I thought then that the argument was well-known.[1] Yet I could not find it presented in print, so I had to reconstruct it for myself. I showed my handout to Paul Teller; he presented the argument, with my permission and with full acknowledgement, in his article 'Conditionalization and Observation'.[2] Teller's article has become the standard source for the argument. But it seems to leave a question in some readers' minds: why does the argument call for conditionalizing on the subject's total increment of experiential evidence, no more and no less? Since my handout had addressed just that question, I decided there was some reason to publish it after all. Apart from a little editing to simplify notation, it appears here in its original form.

The diachronic Dutch book argument can be broken into two halves. Consider a conditional bet: that is, a bet that will be null and void unless its condition is met. We note, first, that the conditional

1 Hilary Putnam alludes to, but does not state, a diachronic Dutch book argument in his 'Probability and Confirmation' in *Philosophy of Science Today*, ed. by Sidney Morgenbesser (Basic Books, 1967), p. 113. He says that if one follows a certain learning rule, it can be shown 'that even if one's bets at any one time are coherent, one's total betting strategy through time will not be coherent'.

2 *Synthese* 26 (1973), pp. 218–258.

bet is equivalent in its outcome, come what may, to a certain pair of unconditional bets. We note, second, that the conditional bet is also equivalent in its outcome, come what may, to a certain contingency plan whereby one's future betting transactions are made to depend on the arrival of new evidence. The first equivalence yields a well-known synchronic argument relating the prices of conditional and unconditional bets. The second equivalence yields a diachronic argument relating the present prices of conditional bets to the future prices, after various increments of evidence, of unconditional bets. We can stitch both halves together and leave the conditional bet unmentioned; and that is the argument presented here.

Richard Jeffrey has suggested that we should respond to experiential evidence not by conditionalizing, but rather by a less extreme redistribution of degrees of belief.[3] Despite appearances, I do not disagree. He and I are considering different cases. My advice is addressed to a severely idealized, superhuman subject who runs no risk of mistaking his evidence, and who therefore can only lose if he hedges against that risk. Jeffrey's advice is addressed to a less idealized, fallible subject who has no business heeding counsels of perfection that he is unable to follow.

Similarly, it seems that we should sometimes respond to conceptual discoveries by revising our beliefs. If first you divide your belief between hypotheses H_1, H_2, H_3, and 'none of the above', and then you discover that 'none of the above' includes a hitherto unnoticed H_4 that is far nicer than the other three, you would be wise to shift some of your belief to H_4, even though you would not be conditionalizing on experiential evidence. Our ideal subject, who never changes his belief except by conditionalizing, will never do that. Is he pig-headed? No – being ideal, he has left no conceptual discoveries unmade. He made them all in his cradle. So he has no occasion to respond to new conceptual discoveries. But we, who are not so smart, would be unwise to emulate him. Some of our departures from ideal rationality are just what we need to compensate for other departures.

Note also that the point of any Dutch book argument is not that it

3 Richard C. Jeffrey, *The Logic of Decision* (McGraw-Hill, 1965; University of Chicago Press, 1983), Chapter 11.

would be imprudent to run the risk that some sneaky Dutchman will come and drain your pockets. After all, there aren't so many sneaky Dutchmen around; and anyway, if ever you see one coming, you can refuse to do business with him. Rather, the point is that if you are vulnerable to a Dutch book, whether synchronic or diachronic, that means that you have two contradictory opinions about the expected value of the very same transaction. To hold contradictory opinions may or may not be risky, but it is in any case irrational.

* * *

Suppose that at time 0, you have a coherent belief function M. Let $E_1, \ldots, E_n$ be mutually exclusive and jointly exhaustive propositions that specify, in full detail, all the alternative courses of experience you might undergo between time 0 and time 1. For each i from 1 to n, let M_i be the belief function you would have at time 1 if you had the experience specified by E_i – that is, if E_i were the true one of $E_1, \ldots, E_n$. You would *conditionalize* if, for any proposition P (in the domain of M),

$$M_i(P) = C(P/E_i) =^{df} M(PE_i)/M(E_i)$$

Why would it be irrational to respond to experience in any other way?

Assume that your belief functions both at times 0 and 1 can be measured by your betting behavior, as follows: your degree of belief that P is the price at which you would be willing either to buy or to sell the bet [\$1 if P, 0 otherwise]. Assume also that if any betting transactions are acceptable to you, so are any sums or multiples thereof.

Suppose $M_i(P)$ is less than $C(P/E_i)$. Then I can follow this three-step plan to exploit the fact.

(1) Sell you the two bets

[\$1 if PE_i, \$0 otherwise]
[\$$x$ if not-E_i, \$0 otherwise]

where $x = C(P/E_i)$, for the maximum price you will pay: *viz.* $\$M(PE_i) + \$xM(\text{not-}E_i) = \$C(P/E_i)$.

(2) Wait and see whether E_i is true. (Thus I need to have as much

knowledge as you, but no more; for you also will know by time 1 whether E_i is true.)

(3) If E_i is true, buy from you at time 1 the bet

[\$1 if P, \$0 otherwise]

for the minimum price you will accept: *viz.* $\$M_i(P)$.

If E_i is false, your net loss will be \$0. If E_i is true (regardless of P) your net loss will be $\$C(P/E_i) - \$M_i(P)$, which by hypothesis is positive. As a result of your failure to conditionalize, I can inflict on you a risk of loss uncompensated by any chance of gain; and I can do this without at any point using knowledge that you do not have.

Likewise if $M_i(P)$ is greater than $C(P/E_i)$ I can exploit that by the opposite plan: buy at step (1), sell at step (3).

If you can be thus exploited you are irrational; so you are rational only if you conditionalize.

Why doesn't a parallel argument work for *any* set $D_1, \ldots, D_n$ of mutually exclusive and jointly exhaustive propositions, showing that your belief function ought to evolve by conditionalization on the true one of *this* set? If $M_j(P)$ is less than $C(P/D_j)$, why can't I take advantage of this?

(1) Suppose D_j is wholly contained in (implies) some E_i, but $D_j \neq E_i$. Then to carry out my plan of exploitation, I must learn that D_j while you learn only that E_i. It proves nothing derogatory about your rationality that I can exploit you by taking advantage of my greater knowledge.

(2) If $D_j = E_i$, I can take advantage of you, but this adds nothing to the argument that you should conditionalize on the true one of $E_1, \ldots, E_n$.

(3) Otherwise D_j overlaps two or more distinct E's; thus you can distinguish two or more ways for D_j to come true, and it is not legitimate to assume that there is a *unique* new belief function M_j that you will end up with if D_j is true. We should consider separately the various belief functions determined by the different

distinguishable ways for D_j to be true; we thus revert to cases (1) and (2).

It has been pointed out[4] that if you fail to conditionalize, I still have no safe strategy for exploiting you unless I *know* in advance what you do instead of conditionalizing. That is: I must know whether $M_i(P)$ is less than or greater than $C(P/E_i)$. But suppose you don't know this yourself. Then I can reliably exploit you only with the aid of superior knowledge, which establishes nothing derogatory about your rationality. – Granted. But I reply that if you can't tell in advance how your beliefs would be modified by a certain course of experience, that also is a kind – a different kind – of irrationality on your part.

4 By D. Kaplan, a student at Princeton in 1972; and by Gilbert Harman.