Undercutting defeat via reference properties of differing arity: a reply to Pust

In a recent article, Pust (forthcoming) argued that direct inference cannot be used to resolve the well known Sleeping Beauty problem (Elga 2000). Pust's article was written in response to an article authored by a group of philosophers, hereafter referred to as 'TMOS' for 'The Members of the Oscar Seminar' (Seminar 2008). TMOS claim that Sleeping Beauty may, during certain waking moments of a Sleeping Beauty scenario, use direct inference to justifiably conclude that the probability is 1/3 that a certain coin (flipped as an element of the scenario) came up heads. The problem with this claim, according to Pust, is that the direct inference to the 1/3-conclusion is rebutted (and thereby defeated) by another equally good direct inference to the conclusion that the probability of heads is 1/2.

I will argue here that Pust's proposed direct inference to the 1/2-conclusion does not have the same standing as the direct inference to the 1/3-conclusion. Rather Pust's proposed direct inference is defeated, because it is based on an incomplete assessment of Sleeping Beauty's relevant evidence. The consequence is that Sleeping Beauty should disregard Pust's proposed direct inference, and accept the direct inference to the 1/3-conclusion.

Typical Bayesian accounts of credence formation focus on *single-case probabilities*. Single-case probabilities attach to propositions, and will be indicated here by the operator 'PROB', as in: PROB(P) and PROB($P|Q$). In contrast to single-case probabilities, *indefinite probabilities* attach to open formulas, and will be indicated by the operator 'prob', as in: prob($\mathrm{T}(x)|\mathrm{R}(x)$). Within the expression 'prob($\mathrm{T}(x)|\mathrm{R}(x)$)', the open formulas '$\mathrm{T}(x)$' and '$\mathrm{R}(x)$' may be thought of as designating the properties T and R, so that 'prob($\mathrm{T}(x)|\mathrm{R}(x)$)' designates the probability of a generic R being a T (or the probability that a generic satisfier of $\mathrm{R}(x)$ satisfies $\mathrm{T}(x)$). Similarly, within the expression 'prob($\mathrm{T}(x)|\mathrm{R}(x,y)$)', we may regard the open formula '$\mathrm{R}(x,y)$' as designating the two-place property R, and regard 'prob($\mathrm{T}(x)|\mathrm{R}(x,y)$)' as designating the probability that the first element of a generic pair satisfying $\mathrm{R}(x,y)$ satisfies $\mathrm{T}(x)$. Regardless of arity, I will call the property associated with the right side of an indefinite probability its 'reference property', while the property associated with the left side is called its 'target property'.

Theories of direct inference codify the conditions under which one is justified in reasoning from indefinite probabilities to single-case probabilities. In circumstances where all of the relevant reference and target properties are unary, standard accounts of direct inference prescribe that one set PROB($\mathrm{T}(\mathrm{c})$) to be prob($\mathrm{T}(x)|\mathrm{R}(x)$), in the case where R is the logically strongest reference property such that one knows that R(c) and one knows the value of prob(T(*x*)|R(*x*)) (Venn 1866; Reichenbach 1949).[1] Pust calls the present prescription ‘Reichenbach’s principle’. In general, where n ≥ m, I will say that an n-place reference property *R′* is logically stronger than an m-place reference property *R* if and only if it is a logical truth that $\forall x_1,\ldots,x_n$: $R'(x_1,\ldots,x_n) \supset R(x_1,\ldots,x_m)$. In those cases where a candidate direct inference is in fact defeated because the reference property of another direct inference is logically stronger, I will say (following Pollock) that the candidate direct inference is subject to *undercutting defeat*.

In making their case regarding Sleeping Beauty, TMOS adopt the following notation (where *x*, *t*, and *s* are variables): ‘H(*x*)’ means ‘*x* lands heads’, ‘Toss(*x*,*s*)’ means ‘*x* is the coin toss involved in *s*’, ‘B(*t*,*s*)’ means ‘*s* is a Sleeping Beauty scenario and *t* a time during *s*’, and ‘W(*t*,*s*)’ means ‘Sleeping Beauty awoke in the scenario *s* sometime during the interval Δ (relative to *t*) and did not remember any previous awakening during *s*’. Using this notation, TMOS argue that Sleeping Beauty is justified in accepting the following indefinite probabilities:

(1) $\text{prob}(H(x)|B(t,s)\wedge \text{Toss}(x,s)) = 1/2$

(2) $\text{prob}(H(x)|W(t,s)\wedge B(t,s)\wedge \text{Toss}(x,s)) = 1/3$

Where σ is the particular Sleeping Beauty scenario in which Sleeping Beauty finds herself, and τ is the coin toss involved in that scenario, TMOS note that, on Sunday before she goes to sleep, Sleeping Beauty knows $B(\text{now},\sigma)\wedge \text{Toss}(\tau,\sigma)$. So on Sunday before she goes to sleep, Sleeping Beauty may conclude, by direct inference using (1), that $\text{PROB}(H(\tau)) = 1/2$. When she is awoken later, she comes to know $W(\text{now},\sigma)\wedge B(\text{now},\sigma)\wedge \text{Toss}(\tau,\sigma)$. Given this new information, Sleeping Beauty may make a direct inference, using (2), to the conclusion that $\text{PROB}(H(\tau)) = 1/3$. At that time, the direct inference based on (1) is subject to undercutting defeat, since (2) involves a logically stronger reference property than (1).

[1] Many recent accounts of direct inference accept variants of the prescription, but (i) are generalized to accommodate interval-valued indefinite probability statements, and (ii) incorporate the observation that the preference for logically stronger reference properties must be tempered, in some cases, in order to deal with the so called ‘projectability’ problems associated with direct inference (Kyburg 1974; Bacchus 1990; Pollock 1990; Kyburg and Teng 2001).

In the face of the preceding approach to the Sleeping Beauty problem, Pust maintains that, in addition to (1) and (2), Sleeping Beauty is justified in accepting the following indefinite probability:

(3) prob(H(*x*)|Toss(*x*,*s*)) = 1/2

The problem, then, according to Pust, is that (3) provides a basis for a direct inference to the conclusion that PROB(H(τ)) = 1/2, and this direct inference is not subject to undercutting defeat by appeal to (2). The result, according to Pust, is that the pair of direct inferences, based on (2) and (3), are mutually defeating, with the consequence that Sleeping Beauty is unjustified in drawing any conclusion about the value of PROB(H(τ)), by direct inference.

The linchpin of Pust's position is the argument for the claim that direct inference based on (3) is not subject to undercutting defeat. Pust's argument proceeds from a description of what he regards as the intuition behind Reichenbach's principle: 'Its basis is the intuition that we should take account of all the properties we know a given object to have in arriving at our credence that the object has some other property.' Given this intuition, Pust immediately proposes a generalization of Reichenbach's principle that is meant to preclude a preference for direct inference based on (2) over direct inference based on (3): 'the generalized version of Reichenbach's principle requires us to base our direct inference to the conclusion that a given *n*-tuple possesses some consequent property (in this case, the property of having its coin toss member land heads) on the logically strongest reference property we know *that very n*-tuple to possess.' Pust's generalized version of Reichenbach's principle requires that Sleeping Beauty use (3) in making a direct inference about ⟨τ,σ⟩, and requires that she use (2) in making a direct inference about ⟨now,σ,τ⟩. But the principle provides no means for arbitrating between the two inferences. Because his generalization of Reichenbach's principle fails to license a preference for direct inference based on (2) over (3), Pust concludes that no preference is licensed: 'This generalized version of Reichenbach's principle still implies that (2) trumps (1), just as TMOS claim, but because (3) and (2) concern property possession by *n*-tuples of different *n*, neither trumps the other as a basis for direct inference.'

Pust's argument is a bit quick. The argument leaps from the claim that some true principle fails to license a preference for one inference over another to the claim that no preference among the two inferences is licensed. To make his case, Pust would have to argue that there is no plausible principle that licenses a preference for direct inference based on (2) over (3). Pust does not argue for that claim, nor does he evaluate existing accounts of direct inference that do license a preference for direct inferences based on logically stronger reference properties, regardless of arity (Pollock 1990; Kyburg and Teng 2001: 216).[2] All that Pust provides is a proposal concerning the intuitive basis of Reichenbach's principle, and a generalization of Reichenbach's principle which is supposed to share that intuitive basis.

The intuition behind Reichenbach's principle is simply that we should prefer direct inferences based on reference properties that incorporate more of the things we know concerning the objects about which we wish to make probability judgments. This intuition *does* support the conclusion that direct inference based on (2) should be preferred to direct inference based on (3). Both (2) and (3) have the potential to bear on Sleeping Beauty's judgments concerning the probability of H(τ). In arbitrating between direct inferences based on (2) and (3), it is decisive that the reference property of (2) incorporates more of the things that Sleeping Beauty knows about τ. In particular, while (3) only incorporates Sleeping Beauty's knowledge that τ is an element of a pair $\langle\tau,\sigma\rangle$, where $\langle\tau,\sigma\rangle$ satisfies Toss(x,s), (2) incorporates that knowledge, as well as her knowledge that τ is an element of a triple $\langle \text{now},\sigma,\tau\rangle$, where $\langle \text{now},\sigma,\tau\rangle$ satisfies $W(t,s)\wedge B(t,s)\wedge \text{Toss}(x,s)$.

Examples that are far less controversial than the case of Sleeping Beauty also demonstrate that Pust's criteria for arbitrating between direct inferences are too narrow. For instance, suppose one knows that there are two films playing at a local cinema, the Loft, on March 1, 2011: a comedy and a documentary. Suppose one also knows that one hundred persons went to the Loft cinema on March 1, 2011, with ninety attending the comedy, and ten attending the documentary. Let 'S(y)' mean 'y was a film *shown* at the Loft cinema on March 1, 2011', 'W(x,y)' mean 'x (a person) *went* to y', and 'D(y)' mean 'y is a *documentary*'. In that case, it is reasonable to accept the following indefinite probabilities:[3]

(4) $\text{prob}(D(y)|S(y)) = 1/2$

[2] Although Pollock does not discuss the point, his principle *DI* (1990: 190) entails a preference for logically stronger reference properties of higher arity, via applications of his principle *IND* (1990: 46).

[3] It is reasonable to accept (4) and (5) (cf. Venn 1866; Reichenbach 1949; Kyburg 1974; Bacchus 1990; Kyburg and Teng 2001), or at least reason by direct inference in accordance with (4) and (5) (cf. Pollock 1990: 70), because the values of (4) and (5) correspond to the known relative frequencies. The relative frequency of objects satisfying S(y) that satisfy D(y) is 1/2, and the relative frequency of pairs satisfying $S(y)\wedge W(x,y)$ whose y element satisfies D(y) is 1/10.

(5) $\text{prob}(D(y)|S(y)\wedge W(x,y)) = 1/10$

Now suppose we come to know of a person named 'Pat' who attended a film at the Loft cinema on March 1, 2011. Further suppose we have no evidence, aside from what has already been described, bearing on whether Pat went to the comedy or the documentary (so that 'Pat' is for us a perfectly nondescript name relative to the predisposition to attend comedies and documentaries). In that case, we could still introduce the name 'f_P' to refer to the film that Pat attended, and attempt to judge the likelihood that f_P was a documentary. Given our information, we should judge it unlikely that f_P is a documentary, since only 10% of the people who attended a film at the Loft on March 1, 2011, attended a documentary. Indeed, under the described circumstances, it is obvious that we should use (5), and our knowledge that $S(f_P)\wedge W(\text{Pat},f_P)$, to make a direct inference to the conclusion that $\text{PROB}(D(f_P)) = 1/10$. It is also obvious that direct inference based on (4) (and our knowledge that $S(f_P)$), to the conclusion that $\text{PROB}(D(f_P)) = 1/2$, is subject to undercutting defeat. But by Pust's criteria, there would be no preference among the direct inferences based on (4) and (5), since their reference properties are of different arity.[4]

In order to get the right answer in the preceding case, and in others cases (including the case of Sleeping Beauty), we must adopt an account of direct inference that licenses a preference for direct inferences based on logically stronger reference properties, regardless of arity.[5]

References

Bacchus, F. 1990. *Representing and Reasoning with Probabilistic Knowledge*. Cambridge, Massachusetts: The MIT Press.

Elga, A. 2000. Self-locating belief and the Sleeping Beauty problem. *Analysis* 60: 143-47.

Kyburg, H. 1974. *The Logical Foundations of Statistical Inference*. Dordrecht: Reidel.

Kyburg, H., and Teng, C. 2001. *Uncertain Inference*. Cambridge: Cambridge University Press.

Pollock, J. 1990. *Nomic Probability and the Foundations of Induction*. New York, New York: Oxford University Press.

[4] Note that there are no available indefinite probability statements over unary reference properties that can be used to make an undefeated direct inference to $\text{PROB}(D(f_P)) = 1/10$. For example, $\text{prob}(D(y)|(\exists x)[S(y)\wedge W(x,y)]) = 1/2$.

[5] For very helpful comments on earlier drafts of this paper, I thank Terry Horgan and Joel Pust.

Pust, J. *Forthcoming*. Sleeping Beauty and direct inference. *Analysis*.

Reichenbach, H. 1949. *A Theory of Probability*. Berkeley, California: University of California Press.

Seminar, O. 2008. An objectivist argument for thirdism. *Analysis* 68: 149-55.

Venn, J. 1866. *The Logic of Chance.* New York, New York: Chelsea Publishing Company.