Luminous Dials for Williamson's Clocks

Kevin T. Kelly
Hanti Lin*

April 27, 2011

Abstract

1 Introduction

"Luminosity" with respect to knowledge means that whenever one has knowledge, one is in a position to know that one has knowledge. Timothy Williamson has a well-known, sorites-like argument, based on the safety requirement for knowledge, for the surprising conclusion that we do not know what we know in ordinary perceptual circumstances (2000). Safety is the requirement that one's true belief could not easily have been false. More recently, in a pair of intriguing examples based on numberless clock faces (2010a, b), Williamson has attempted to extend his position by providing examples in which one knows but one's degree of belief that one knows is extremely small, not due to aphasia or ignorance, but due to the very semantics of knowledge. A crucial feature of Williamson's examples is that the sense of "could easily" is interpreted in terms of We respond that, in cases of visual acuity, the relevant sense of "could easily have been false" pertains to "could easily have been produced in error", rather than "are actually false in nearby worlds". We show that, under the former interpretation, full luminosity is possible and that Williamson's Sorites-like argument is necessarily unsound, even though each premise is true with high objective chance. Thus, contrary to Williamson's insistence, the argument is an instance of the lottery paradox.

*This work was supported generously by the National Science Foundation under grant 0750681. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

2 Williamson's Sorites-Like Argument Against KK

Let t be the height of a tree in inches and let the subject be near-sighted and be located a block away from the tree. Consider the situation in which $t = 666$. Let p_i state that $t = i$. Let n be a large quantity (e.g., 666). Williamson's argument is the following:

1. premise: $K\neg p_0$;
2. premise: $\neg K\neg p_n$;
3. premise: K is closed under logical entailment;
4. premise: $\bigwedge_{i=0}^{n-1} K(p_{i+1} \to \neg K\neg p_i)$;
5. conclusion: $\bigwedge_{i=0}^{n-1} K\neg p_i \to KK\neg p_i$.

The reasoning is valid. Suppose for reductio that $\bigwedge_{i=0}^{n-1} K\neg p_i \to KK\neg p_i$ holds. Williamson proves the following lemma:

$$\bigwedge_{i=0}^{n-1} K\neg p_i \to K\neg p_{i+1}. \tag{1}$$

Suppose that $K\neg p_i$. Then by the reductio hypothesis $KK\neg p_i$. Note that $K\neg p_i$ and $p_{i+1} \to \neg K\neg p_i$ entail that $\neg p_{i+1}$. So since both premises are known (by the lemma hypothesis and the fourth premise), it follows from closure of K under entailment (the third premise) that $K\neg p_{i+1}$. Now we can reason as follows, using the first premise and the lemma:

$$\begin{array}{l} K\neg p_0; \\ K\neg p_0 \to K\neg p_1; \\ K\neg p_1; \\ K\neg p_1 \to K\neg p_2; \\ \vdots \\ K\neg p_n. \end{array}$$

The last line contradicts the second premise. So the argument's conclusion follows by reductio ad absurdum.

What about soundness? The first premise seems indubitable—even blurry vision can detect the green mass of the tree, which must occupy some nonzero height. We also grant the second premise, because $\neg p_n$ is false and, therefore, cannot be known. The relevant instances of closure of K under entailment do not seem troublesome, especially after Williamson reminds us of the entailments. The fourth premise is the crucial one, but it also seems very plausible—blurry vision can't resolve inches from a block away. So the KK principle must fail for at least one height between 0 and

$n-1$—the argument doesn't say which. Williamson would prefer to conclude that it always fails.

A more detailed argument g

Let $t : W \to \mathbb{R}$. Define:

$$R(w, w') \;\Leftrightarrow\; w' \text{ is close to } w; \tag{2}$$
$$w \models S\phi \;\Leftrightarrow\; (\forall w' \in R(w))\; w' \models \phi; \tag{3}$$
$$w \models t = r \;\Leftrightarrow\; t(w) = r. \tag{4}$$

Interpret $S\phi$ as "it is safely the case that ϕ". Define "it is easily the case that ϕ as the dual of safety:

$$E\phi \;\Leftrightarrow\; \neg S \neg \phi. \tag{5}$$

Think of p as a statement whose truth depends only on the value of t. Williamson assumes the following theses:

1. $\models K\phi \to B\phi$;

2. $\models K\phi \to S\neg(B\phi \wedge \neg\phi)$;

3. $\models \Box(t = r \to p) \vee \Box(t = n \to \neg p)$, for all r;

4. $\models \Box(Bp \wedge t = u \wedge |u - v| < c \to E(t = v \wedge Bp))$.

The *margin for error principle* follows.

Proposition 1. $|t(w) - t(w')| < c \wedge w \models Kp \Rightarrow w' \models p$.

Thus:

Proposition 2. $|t(w) - t(w')| < c \Rightarrow w \models p \Leftrightarrow w' \models p$.

Thus, if all values of t are possible, p is necessarily true or necessarily false.

The point of Williamson's later efforts (2011a, b) is to extend the preceding argument to show that the KK principle essentially does always fail even in such cases of visual acuity and that, moreover, it fails in a spectacular way: the subject even finds it very *probable* that she does not know what she knows by looking. Even more strikingly, the KK principle does not fail for familiar reasons such as ignorance of one's own beliefs or one's own acuity, but due entirely and ineluctably to the very *logic* of knowledge.

3 Basic Epistemic Logic

Williamson's new arguments are framed within standard, modal, epistemic logic, whose elements we quickly review. A *Kripke frame* for epistemic logic is just a pair (W, R), where W is a set of "possible worlds" and R is a binary, reflexive "accessibility relation" over W. Propositions are just subsets of W. K is an operation on propositions. The intended interpretation is that $K\phi$ is the proposition "S knows that ϕ". Formally, the truth conditions for $K\phi$ are as follows:

$$w \in K\phi \quad \Longleftrightarrow \quad (\forall w' \in W)\ R(w, w') \Rightarrow w' \in \phi. \tag{6}$$

Clearly, much depends on the choice of R. If we had some antecedent idea of what "epistemic world accessibility" is, then definition (6) would be an explanatory *analysis* of knowledge in terms of R. But for obvious candidates, that is a terrible idea. For example, suppose that $R(w, w')$ means "for all S has been infallibly informed in w, the actual world might be w'". Then $K\phi$ would imply that ϕ is a deductive consequence of infallible information—a quick recipe for radical skepticism. Williamson reverses the explanatory order by defining R in terms of K, so that $R(w, v)$ says only that for all S *knows* in world w, the actual world is v. By formulating his clock examples in Kripke semantics, Williamson can hold his cards close regarding what knowledge *is*, leaving the critic to fill in and to defend the missing details, such as what S believes, how visual perception works, and what constitutes visual justification.

4 Knowledge States

Even on Williamson's non-committal interpretation, Kripke semantics is not entirely vacuous regarding the nature of knowledge. The assumption that any sharp-edged proposition is known in the case of perception is open to question. Perception could guide actions even if it resulted only in a posterior probability distribution over positions of things. Propositions could be a rough-and-ready way to communicate and reason with more nuanced, underlying, degrees of belief.[1] In that case, the assumption that knowledge is propositional is already a *faux pas*. And even if it is conceded that some propositions come to be known by perception, epistemic logic commits one to the unrealistic view that knowledge is closed under logical entailment.[2] It follows that we already know all of mathematics and the results of all lengthy computations. Presumably, it remains only to "recollect" them via "stimulation" by the actual proofs

[1] For a principled account of how that might go without inviting the lotter paradox, cf. (Lin and Kelly 2011).

[2] For suppose that $w \in K\phi$ and ϕ logically entails ψ. Then each world accessible from w makes ϕ true and each such world also makes ψ true, so $w \in K\psi$.

and computations. We concede these assumptions, however, because our principal concern lies elsewhere.[3]

Given the deductive closure condition, Williamson represents the overall knowledge state of S as the strongest proposition known by S. Define:

$$R(w) \quad = \quad \{w' \in W : R^n(w, w')\}. \tag{7}$$

The truth conditions for $K^n\phi$ can be re-expressed succinctly as follows:

$$w \in K\phi \quad \Longleftrightarrow \quad R(w) \subseteq \phi. \tag{8}$$

In the case in which $\phi = R(w)$, we have:

$$w' \in KR(w) \quad \Longleftrightarrow \quad R(w') \subseteq R(w). \tag{9}$$

Setting $\phi = KR(w)$ yields:

$$w \in KKR(w) \quad \Longleftrightarrow \quad R(w) \subseteq KR(w) \tag{10}$$

$$\Longleftrightarrow \quad (\forall w' \in R(w))\ w' \in KR(w). \tag{11}$$

Now $w' \in KR(w)$ can reexpressed via (9):

$$w \in KKR(w) \quad \Longleftrightarrow \quad (\forall w' \in R(w))\ R(w') \subseteq R(w). \tag{12}$$

This formulation makes it apparent that there are *degrees* of failure of the KK principle for $R(w)$ in world w. Of course, $w \in R(w)$ because $R(w)$ is known in w. If $w \in KKR(w)$, $R(w') \subseteq R(w)$ holds for each w' in $R(w)$. A *maximal* failure of $w \in KKR(w)$ occurs when w is the only world w' in $R(w)$ for which $R(w') \not\subseteq R(w)$. A *minimal* failure occurs when there is only one world w' in $R(w)$ such that $R(w') \not\subseteq R(w)$. And then there are all the cases in between.

In order to measure degrees of failure of the KK principle, Williamson calculates the fraction:

$$p(KKR(w) \mid R(w)) \quad = \quad \frac{|\{w' \in W : R(w') \subseteq R(w)\}|}{|R(w)|}. \tag{13}$$

Since $R(w)$ is the knowledge state of S, the quantity $p(KKR(w) \mid R(w))$ can be viewed as the "objective" probability of $KKR(w)$, given everything that S knows. Williamson refers to this as the *evidential* probability of $KKR(w)$ in w. We will not venture to debate the aptness of his terminology here. Better, for present purposes simply to

[3] For systematic developments of the strong analogies between empirical and formal inquiry, cf. (Kelly ***, ***, ***).

refer to think of $p(KKR(w) \mid R(w))$ as a measure of the *degree* to which the KK principle holds for S with respect to the knowledge state $R(w)$ of S. In particular, $p(KKR(w) \mid R(w)) = 1$ if $w \in KKR(w)$ and $p(KKR(w) \mid R(w)) = 1/|R(w)|$ in the worst cases. Therefore, as $|R(w)|$ goes to infinity, the worst case value of

$$p(KKR(w) \mid R(w))$$

approaches zero. Williamson's goal is to show that the worst case happens routinely in ordinary cases of visual perception.

5 Safety

In Williamson (2000), the crucial premise of the sorites-like argument is motivated in terms of *safety from error*. Let $S\phi$ abbreviate that "it is safely the case that ϕ".

6 Williamson's First Clock

In (2010a), Williamson invites the reader to consider a modernist clock dial with a single pointer and a blank face (fig. 1.a). The hand position is digitized, so it can point at one of N discrete positions $0, 1, \ldots, N-1$ and let T be the set of all such positions. Let $\rho(t, t')$ denote the distance between positions t and t' on the clock face, measured in terms of the nearest number of discrete steps between t and t' (fig. 1.b). In the example, the clock is not moving (maybe the battery is missing) and S is looking at the clock from some distance, so its hand is stuck in one of the N positions.

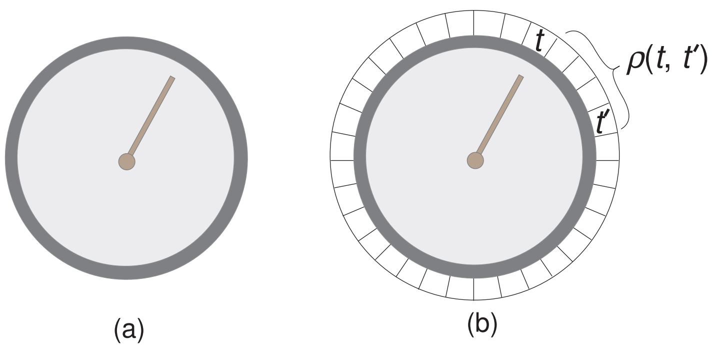

Figure 1: Williamson's modernist clock

For ease of presentation, Williamson takes the set W_1 of possible worlds to be the N possible, discrete positions of the pointer:

$$W_1 = T. \tag{14}$$

That seems reasonable enough—it is a routine practice to exclude irrelevant clutter from models. According to Williamson, it is plausible to assume that S can only discern that the pointer is within $d \leq N$ units, left or right, of the true position t, where d depends on the visual acuity of S, her distance from the clock, clarity of the intervening air, lighting, etc. (figure 2.a). Formally:

$$R_{1,d}(t, t') \iff \rho(t, t') \leq d, \tag{15}$$

That also seems plausible—exact precision can't be expected at every distance and it

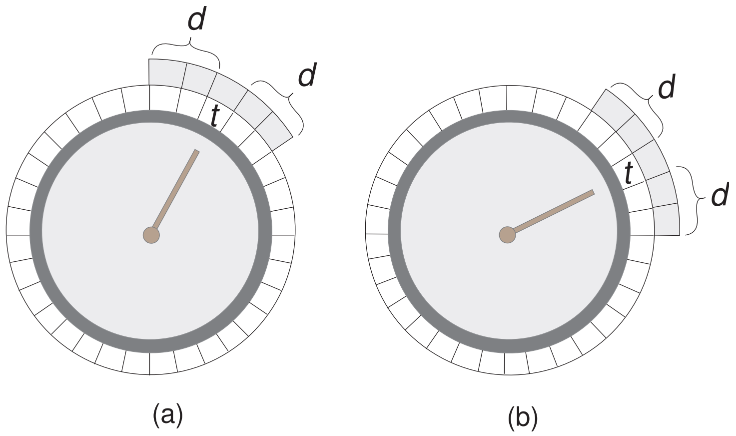

Figure 2: Williamson's perfectly centered known intervals

is very implausible that there is a particular distance at which perfect resolution stops and perfect ignorance starts.

The KK defender is already trapped, for the only world w' such that $R(w') \subseteq R(w)$ is w, itself. Hence, the model realizes the worst case failure of the KK principle, so as N and $d \leq N$ go to infinity, the evidential probability $p(KKR(w) \mid R(w))$ approaches zero. The result does not rely on holograms, vats, or the usual philosophical legerdemain. Nor does it hinge on a banality such as assuming that S has aphasia and does not even believe that she believes $R_{1,d}(w)$—belief is not even mentioned in the model. And the underlying logical framework adds almost no content to the example aside from logical omniscience, which does not seem to do any heavy lifting in the example. It is almost as though Williamson's skepticism is pulled from thin air—or pure logic. One can hardly resist taking a closer look.

7 Improbability of Williamson's First Clock

Williamson's motivation for modeling perceptual knowledge with intervals is not entirely clear, but we assume that it is related to his discussion of safety.

Williamson does not expand on the underlying reasons for that fact. One standard reason for reporting measurements as intervals rather than as point values is that myriad, independent causes of error collaborate to generate off-center measurements, with a low chance of producing wildly inaccurate measurements. Therefore, if one reports a sufficiently large interval around the measured value one obtains, there is a high chance—called the *confidence level*—that the interval contains the true value. However, on that story, the chance is 1 that the interval so produced is centered perfectly. In the clock example, readings are discrete, so the chance of obtaining a perfectly centered interval is non-zero, but it can be driven arbitrarily low by reducing the distance between the discrete readings. So after one generates such a perceptual interval, one hardly knows that it is perfectly centered on the true reading (fig. 2. a,b). Far more plausibly, one knows that it is not—but we need not insist on that point.

Perhaps Williamson would reject that whole story—some sort of mental vagueness aside from visual acuity is responsible for dilating the interval. But that is still implausible. In Williamson's model, S knows that her perceptual interval is perfectly centered. It is hard to see how she could know any such thing unless either (i) she made a lucky guess that Williamson counts as knowledge or (ii) her visual cortex is somehow capable of recording perfectly accurate positional information that her higher cognition somehow fumbles.

HERE

Even if S does not know the true observation exactly, her visual system must at some level record the clock reading with perfect accuracy in order to cause perfectly centered intervals.

The perfect centering assumption is also wildly implausible, for even if S does not know more than an interval, her brain must somehow capture perfect positional information at some hidden level and then censor it from the cognitive level pertinent to propositional knowledge. One wonders how S's hum-drum 20/30 eyesight can see better than the best telescopes and why her higher-level cognition throws that miraculously accurate information away.

None of that would matter if Williamson's anti-luminosity conclusion did not depend on the perfect centering assumption, but it depends entirely on that assumption. Because intervals are perfectly centered, the only world in which S's knowledge state is the same as in the actual world is the actual world itself. To put the point another way, S knows that her intervals are perfectly centered, so if she did know exactly what her known interval is, then she could *deduce* the exact clock reading as the center of that interval. But she can't plausibly know the exact clock reading, so she must not know what her known interval is. If she did not know that her known interval is perfectly centered, the absurd conclusion would not follow from her knowing what her known interval is. We conclude that Williamson's first clock model is fanciful and that the strong non-luminosity it implies is a direct consequence of the fancy.

8 Williamson's Second Clock

The preceding difficulties stem from Williamson's well-intentioned pedagogical decision to simplify the presentation of the model by identifying worlds with clock readings. That initial decision made it impossible for him to model ignorance about centering of the knowledge state, since such ignorance requires that multiple possible centers be associated with the same true clock reading. Williamson corrects that fault in a second article (***) that presents a more complex model in which $KKR(w)$ again fails to the maximum degree.

Each possible world is now assumed to be a pair (t, c), where t is the true clock reading and c is the center of the clock interval known by S. Williamson imposes no constraints on such pairs, so the set W_1 of worlds in the model is the set of all pairs of positions:

$$W_2 \quad = \quad T \times T. \tag{16}$$

Williamson proceeds to define the accessibility relation $R_{2,d}$. Since $R_{2,d}(t, c)$ is now a set of pairs, it is convenient to define the *interval known* by S in (t, c) to be:

$$I_{2,d}(t, c) \quad = \quad \{t \in T : (\exists c' \in T)\ (t', c') \in R_{2,d}(t, c)\}. \tag{17}$$

In an impressive show of bravado, Williamson generously concedes to the luminist that S is self-aware enough to know the midpoint of her own known interval:

$$R_{2,d}((t, c), (t', c')) \Rightarrow c = c'. \tag{18}$$

We accept his offer, with the assurance that if S is assumed to know an interval and is also *assumed* to be irremediably clueless what that interval is, we are happy to concede that S is non-luminous—but only under those assumptions. Williamson's second constraint is inherited from his first model, namely, that the interval known by S is never narrower than $2d + 1$:

$$\rho(t, t') \leq d \Rightarrow R_{2,d}((t, c), (t', c)). \tag{19}$$

We grant that assumption as well, assuming that d cannot be reduced without unduly compromising confidence. Finally, of course, knowledge is true, so:

$$R_{2,d}((t, c), (t, c)). \tag{20}$$

Williamson then proposes that the "obvious" way to realize these constraints is to let the width of the interval known by S in (t,c) *dilate* by adding the distance of the interval's midpoint c from the true clock reading t to the fixed distance d so that the total interval width is $2(\rho(t,c)+d)+1$ rather than the fixed width $2d+1$ assumed in Williamson's first model (fig. 3).

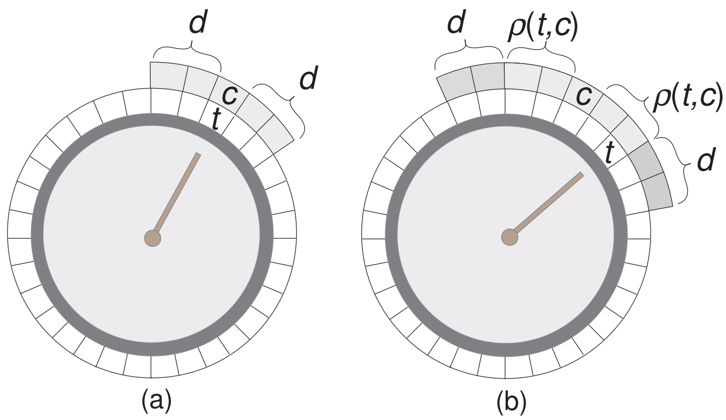

Figure 3: Williamson's dilating confidence intervals

$$R_{2,d}((t,c),(t',c')) \iff c = c' \wedge \rho(t',c) \leq \rho(t,c) + d. \tag{21}$$

It is immediate from the definition that $R_{2,d}$ satisfies the required properties (18, 19, 20).

Now Williamson obtains his intended result—at least in the world (t,t) in which the known interval is exactly centered. Suppose that $t' \neq t$. Then, due to the dilation of $I_{2,d}(t,c)$ as $\rho(t,c)$ increases, $R(t,t') \not\subseteq R(t,t)$. Thus, the unique world (t,c) such that $R(t,c) \subseteq R(t,t)$ is (t,t), itself—the worst case failure of the KK principle for $R(t,t)$.

9 Improbability of Williamson's Second Clock

Williamson is more reticent concerning what happens in his second model in worlds other than (t,t). In fact, his model says that luminosity increases dramatically as the known interval becomes more off-center. For suppose that $\rho(t,c) \geq kd$. Then, $2kd$ of the total of $2kd + 2d + 1$ worlds have clock readings t' that are *closer* to the center c of the interval than t is, so those worlds have narrower known intervals on the same centers, so by condition (9), each such world is in $KR_{2,d}(t,c)$. By letting N and k grow without bound, the evidential probability $p(KKR_{2,d}(t,c) \mid R_{2,d}(t,c))$ approaches 1, so S becomes more luminous without bound as the interval she knows is more off-center and the clock positions become closer together. This sharp and startling implication of the model does not seem to have any basis in ordinary thinking about knowledge and Williamson provides no guidance for interpreting it.

The extreme non-luminosity of the model would still be interesting, in spite of its strangeness, if the non-luminosity followed entirely from the plausible principles (18, 19, 21). But, in fact, it depends entirely on the ad hoc dilation rule that gives rise to the strangeness. Williamson is correct that *some* sort of dilation term is required to ensure principle (refawarax), but the dilation need not occur when the confidence interval of size $2d + 1$ happens to be true. And when that interval is false, the dilation need not extend beyond the true value. In fact, in light of the preceding discussion of confidence intervals, that alternative dilation rule is far more intuitive than Williamson's. In the next section, we present a pair of alternative dilation rules that have an intuitive underlying rationale in terms of confidence intervals.

10 The Clock Illuminated

Suppose that given the ambient lighting, the distance of the clock, the visual acuity of S, etc, S knows that she can achieve a reasonable confidence level by reporting the clock reading within an interval of width $2d + 1$. Suppose that S is also so fortunate as to know that there is no Gettier apparatus or other philosophical monkey business around. That covers just about all the bases—except for truth. So we propose that there is no Williamsonian dilation as long as the confidence interval of width $2d + 1$ produced by S is true. In terms of accessibility (fig. 4.a):

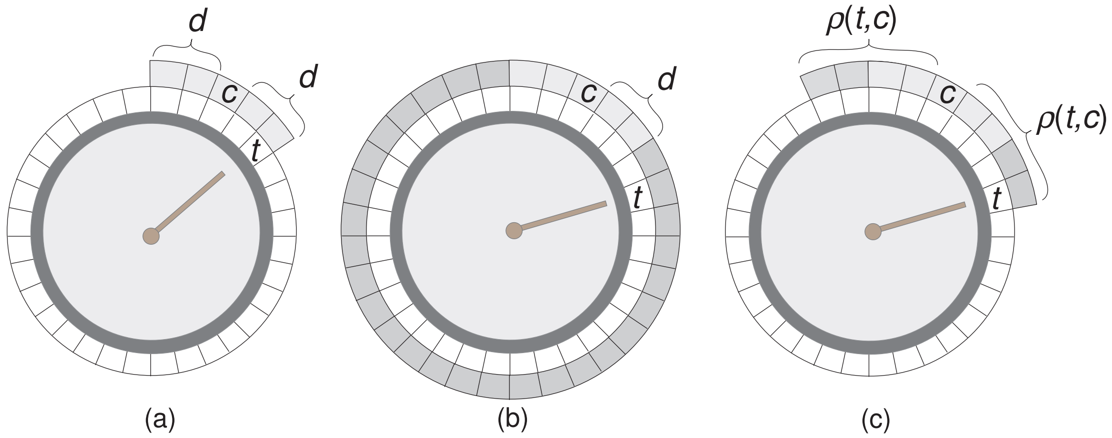

Figure 4: perceptual knowledge based on confidence intervals

$$\rho(t,c) \leq d \quad \Rightarrow \quad (R_{3,d}((t,c),(t',c')) \iff c = c' \wedge \rho(t',c) \leq d)\,. \tag{22}$$

But now that there is no dilation as long as $\rho(t,c) \leq d$, the knowledge state of S no longer depends on the distance $\rho(t,c)$, so $R_{3,d}(t',c) \subseteq R_{3,d}(t,c)$ when $\rho(t,c) \leq d$. So in this case $(t,c) \in KKR_{3,d}(t,c)$, so S is perfectly luminous and Williamson's quest for extreme anti-luminosity comes up empty handed. Since the confidence interval is true with high chance, the chance that S is not perfectly luminous is at worst very small. So Williamson's anti-luminist conclusion is already extremely improbable—in objective chance, not mere opinion. Perhaps that is satisfactory enough, from a luminist point of view.

But let's move on to the case in which the confidence interval of width $2d+1$ is false. What does S know, when her knowledge is grounded on the ability to produce intervals of width $2d+1$ with suitably high confidence and when the interval so produced is false? That is an interesting question. The answer depends on details about S that are not settled in the model. If she believes only the interval she produces and it is false, one might say that she knows *nothing* about the clock's reading—she merely has a false belief (fig. 4.b). But S still knows what the center c of her confidence interval is, so in terms of accessibility the proposal is:

$$\rho(t,c) > d \quad \Rightarrow \quad (R_{3,d}((t,c),(t',c')) \iff c = c')\,. \tag{23}$$

So when $\rho(t,c) > d$, we have that $R_{2,d}(t,c) = T \times \{c\}$. But then it is evident that $R_{3,d}(t',c) \subseteq R_{3,d}(t,c)$, so again the KK principle holds perfectly and S is perfectly luminous.

Alternatively, one could concede to Williamson that dilation is gradual in case the confidence interval of width $2d+1$ is false. The procedure of believing a wider interval on the same center as the interval she produces has higher confidence than the procedure she follows. By deductive closure, she believes larger intervals containing the intervals she believes. She is in a position to know all of this. So she is at least as justified in believing the larger interval as she is in the small interval. On this more optimistic view, her knowledge state is the narrowest true super-interval with the same center as her belief state (fig. 4.c). Formally, let $R_{4,d}$ agree with $R_{3,d}$ when $\rho(t,c) \leq d$ and in the $\rho(t,c) > d$ case define:

$$\rho(t,c) > d \quad \Rightarrow \quad (R_{4,d}((t,c),(t',c')) \iff c = c' \wedge \rho(t',c) \leq \rho(t,c))\,. \tag{24}$$

Thus, in general:

$$R_{4,d}((t,c),(t',c')) \iff c = c' \wedge \rho(t',c) \leq \max(d, \rho(t,c)). \tag{25}$$

Thus, $R_{4,d}$ plausibly substitutes maximization for the summation that occurs in Williamson's definition of R_2. One might object that the alleged knowledge is Gettiered—S's true belief is based on a false belief. In response, her justification is that the process generating larger intervals has even higher confidence, which is not false. Still, it might seem that this story makes knowledge too easy, since there will always be *some* interval S knows no matter how screwed up her observation happens to be. A potential response is that if the disturbance that led to the huge error was "drawn" from the distribution that the bell curve represents, then it is already factored into her justification, which is based on that distribution. Some epistemologists might object that the outlier is in a "different reference class". That is an interesting, material epistemological debate that we need not settle now, since Williamson's anti-luminist argument also fails in this alternative model. For whenever the $\rho(t, c) > d$, the known interval is *maximally* off-center. And when $\rho(t', c) \leq \rho(t, c)$, it follows that the center c is the same and the known interval is no larger, so once again $R_{4,d}(t', c) \subseteq R_{4,d}(t, c)$, the KK principle holds perfectly, and S is perfectly luminous. That is remarkable, given that the width of $R_{4,d}$ depends on the true position t in the case under consideration.

These models describe the clock task more naturally than either of Williamson's models and they are also motivated by a little explanatory story based on standard statistical ideas about measurement accuracy. But S is perfectly luminous in both.

11 Failure of Williamson's Sorites Argument

It seems that Williamson views his clock models as mere icing on the anti-luminosity cake—the heavy lifting is supposed to be done by his celebrated Sorites argument against luminosity. The argument has generated an extensive literature that we will not review here, as our purpose is merely to focus on how the argument fares in our models. We take the liberty to adapt Williamson's argument to the clock setting. Suppose that the clock face is large and that the N positions $1, \ldots, N$ are sufficiently close together to make them impossible to discriminate but the face is wide enough that $N/2$ is easily discriminable from 1 on the opposite side of the face. Let q_i denote the proposition that the true value is exactly i. Some clock reading is true, say position 1. In that case, due to the proximity of the clock readings, we have:

$$\neg K \neg q_2. \tag{26}$$

Given that the clock face is wide, S knows that the clock does not read the diametrically opposed value $N/2$:

$$K \neg q_{N/2}. \tag{27}$$

It seems that S would have to have an unrealistically sharp focus on her own belief state to know exactly where her confidence interval ends. So Williamson assumes that, for each i such that i such that $1 < i \leq N/2$, if S knows that she knows that the true reading is not $i+1$, then she also knows that the true value is not i:

$$KK\neg q_i \quad \rightarrow \quad K\neg q_{i-1}. \tag{28}$$

Now, assume for reductio argument that S is luminous. That implies that for each $i \leq N/2$:

$$K\neg q_i \quad \rightarrow \quad KK\neg q_i. \tag{29}$$

The $N/2$ respective statements of forms (28) and (29) yield, for each $i \leq N$, that:

$$K\neg q_i \quad \rightarrow \quad K\neg q_{i-1}. \tag{30}$$

By (27) and the N instances of (30) one obtains:

$$K\neg q_2, \tag{31}$$

which contradicts (26). So some statement of form (29) is false and, hence, so is luminosity.

We have just seen that S is luminous in models $(W_2, R_{3,d})$ and $(W_2, R_{4,d})$. Since the models also validate Williamson's epistemic logic, it must be that the models make some premise false. Here is a very interesting feature of Williamson's argument—it is *necessarily unsound*, in the sense that some premise or other is false in every possible world, but whichever premise happens to fail in a given world does so with extremely low chance in that world. So S *knows* that the argument is unsound. Nonetheless, in light of the logic of confidence intervals, each individual premise has a very high chance of coming out true! So the usual practice of running down the list of premises and checking them for plausibility results in judging the argument to be sound. The subtlety of the argument's unsoundness helps to explain its persistent interest and also serves as a caution against a piece-meal approach to evaluating the soundness of Sorities-like arguments.

It facilitates our analysis to define the perceptual interval known by S in (t, c) as follows:

$$T_{3,d}(t, c) \quad = \quad \{t' < N : (\exists c' < N)\ R_{3,d}((t, c), (t', c'))\}; \tag{32}$$

Since $T_{3,d}(t, c)$ is an interval over the clock face, $T_{3,d}(t, c)$ has a right (maximum) endpoint if $T_{3,d}(t, c)$ does not cover the entire clock face.

1. Clock reading 1 is the right endpoint of $T_{3,d}(1, c)$ (figure 5.a). Then premise (26) fails, since S knows that the true reading is not $t-1$. But the failure only occurs with very low chance, because the chance is extremely small that the confidence

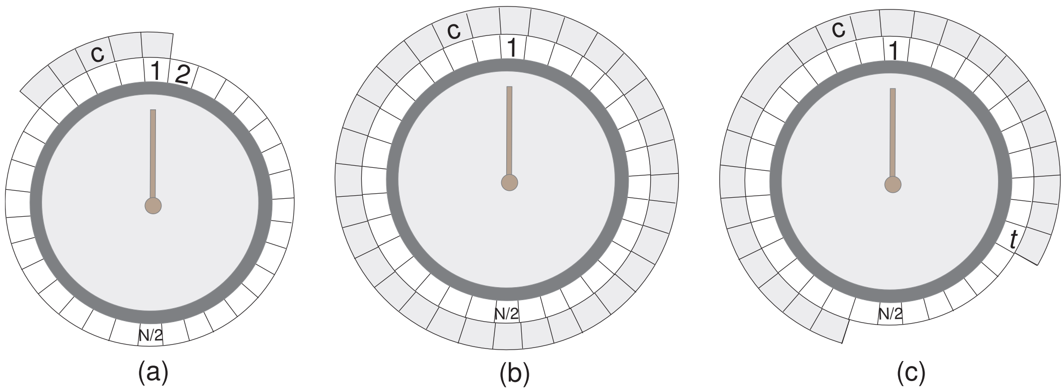

Figure 5: failure of the Sorites argument

interval would be sampled all the way to the left with respect to the true clock reading.[4]

2. The known interval $T_{3,d}(1, c)$ includes $N/2$ (figure 5.b). Then premise (27) fails. But given that the disk of the clock is large, it would be extremely improbable to sample a confidence interval that far away from the true clock reading 1.

3. The known interval $T_{3,d}(1, c)$ does not include $N/2$ and t is not the right endpoint of $T_{3,d}(t, c)$ (figure 5.c). Position 1 is in $T_{3,d}(1, c)$ but $N/2$ is not in $T_{3,d}(1, c)$, so there is a right endpoint t of $T_{3,d}(1, c)$ and t falls properly between 1 and $N/2$. The conditional of form (28) for $i = t + 1$ is false. For recall that KK holds in the model, so since $t + 1$ is outside $T_{3,d}(1, c)$, we have that $\neg q_{t+1}$ is entailed by $T_{3,d}(1, c)$ and, hence, is known. Since KK holds in the model, $KK\neg q_{t+1}$ is true at $(1, c)$. But t falls inside the interval, so q_t is not known. But again, it is very hard to sample a confidence interval whose endpoint is exactly t, so the chance that the $t + 1$th instance of (28) failing is extremely small.

12 Assumed Ignorance of Centering

Williamson's anti-luminosity is not inevitable—we have just explained its slippery unsoundness in the confidence interval model of perception. But the perfect luminosity of our models is not inevitable either. Our models assume, like Williamson's, that S is aware of the exact radius d and center c of her confidence interval. In light of

[4]It is usually assumed that the sampling distribution in measurement is approximately Gaussian and centered on the true clock reading. Thus, it is extremely improbable to sample an interval whose center is so far away from 1 in this case, under the assumption that the positions are very close together.

the preceding discussion, Williamson may prefer to retract those generous concessions. However, without the concessions, Williamson's models lose their magic. If it is simply *assumed* that S does not know what her knowledge state is, then it is immediately apparent that S does not know what her knowledge state is and, hence, does not know that she knows the strongest proposition she knows. Nobody needed a Sorites argument or fancy logic for that.

Williamson is still in a position to score an interesting point against luminisim, however. The luminist might be happy if slight failures of S to access her known interval of clock readings were to result in large failures of the KK principle, as measured by Williamson's evidential probability that she knows what she knows. In fact, that is the case, as we now demonstrate.

We begin by relaxing S's perfect knowledge of the center c of her own known confidence interval. To keep notation under control, rename c as c_0 and d as d_0. Now introduce new parameters c_1, d_1 such that S has generated a higher-order confidence interval around c_0 whose midpoint is c_1 and whose width is $2d_1 + 1$ (fig. 6.a).

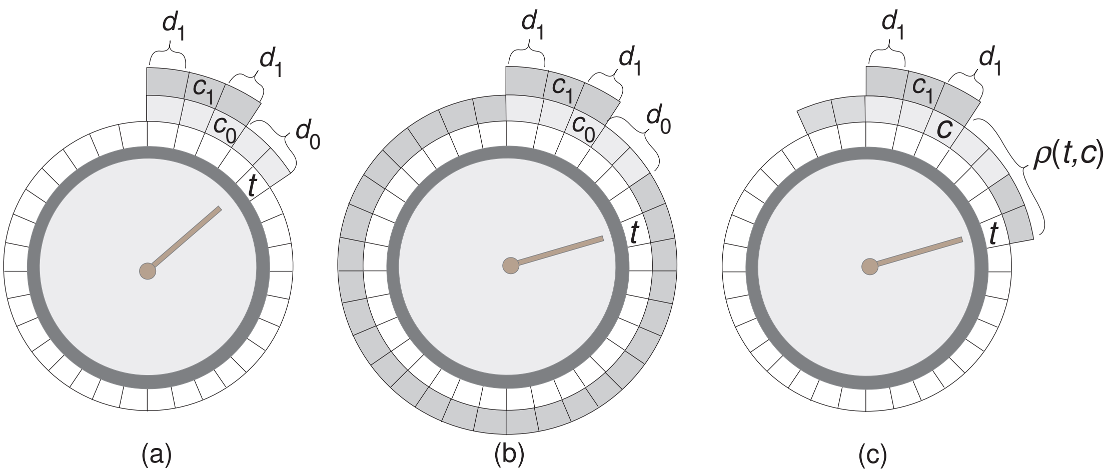

Figure 6: ignorance of centering

$$\rho(t, c) \leq d \quad \Rightarrow \quad (R_{3,d}((t, c), (t', c')) \iff c = c' \wedge \rho(t', c) \leq d) \, . \tag{33}$$

For now, assume that S is perfectly luminous about d_0 and d_1—of course, Williamson might question that as well, but it turns out that he doesn't need to in order to make his point. That leaves the value of c_1 free to vary, so each possible world is now a triple (t, c_0, c_1). Let W_3 denote the set of all such possible worlds. Recall that there are two policies for defining S's knowledge state when her confidence interval over a parameter is false: (i) she knows nothing about the parameter (fig. 6.b) or (ii) she knows the least interval on the same center that contains the true parameter value (fig. 6.c). We develop option (ii) here, leaving the easier case (i) for the reader. Consider world (t, c_0, c_1) with respect to fixed parameters d_0, d_1. The strongest interval for t known by S is the smallest interval centered on c_0 of width $\geq 2d_0 + 1$ that contains t. The strongest interval for c_0 known by S is the smallest interval centered on c_1 of width $\geq 2d_1 + 1$ that contains c_0. And we allow, on Williamson's behalf, the liberal concession that at least c_1 is known perfectly by S:

$$R_{4,d_0,d_1}((t, c_0, c_1), (t', c'_0, c'_1)) \iff \begin{cases} c'_1 &= c_1; \\ \rho(t', c_0) &\leq \max(d_0, \rho(t, c_0)); \\ \rho(c'_0, c_1) &\leq \max(d_1, \rho(c_0, c_1)). \end{cases} \tag{34}$$

Note that:

$$R_{4,d_0,0} = R_{4,d_0}, \tag{35}$$

so $R_{4,d_0,0}$ validates KK. We now proceed to show that the violation of KK can be severe, in Williamson's sense of evidential probability, even when $d_1 = 1$. Let $w = (t, c_0, c_1)$ be an arbitrary world.

Knowledge is a bit more complicated in this model, since it concerns both the interval of clock readings known by S and the (dissociated) interval of possible centers of the former interval that is known by S. Let $T_{4,d_0,d_1}(t, c_0, c_1)$ be the former interval (in world (t, c_0, c_1) and let $C_{4,d_0,d_1}(t, c_0, c_1)$ be the latter.

$$T_{4,d_0,d_1}(t, c_0, c_1) = \{t' < N : (\exists c'_0 < N)\ R_{4,d_0,0}((t, c_0, c_1), (t', c'_0, c_1))\}; \tag{36}$$

$$C_{4,d_0,d_1}(t, c_0, c_1) = \{c'_0 < N : (\exists t < N)\ R_{4,d_0,0}((t, c_0, c_1), (t', c'_0, c_1))\}. \tag{37}$$

It is easier to think of accessibility in terms of these known intervals. Each world (t', c'_0, c'_1) accessible from (t, c_0, c_1) has component c'_0 drawn from $C_{4,d_0,d_1}(t, c_0, c_1)$ and component t' drawn from $T_{4,d_0,d_1}(t, c_0, c_1)$. Component c'_1 is fixed at the value c_1:

Proposition 3.

$$R_{4,d_0,d_1}(t, c_0, c_1) = T_{4,d_0,d_1}(t, c_0, c_1) \times C_{4,d_0,d_1}(t, c_0, c_1) \times \{c_1\}. \tag{38}$$

Proof. The $\subseteq$ inclusion is immediate, by definitions (36) and (37). For inclusion $\supseteq$, suppose that (t', c'_0, c'_1) is not in $R_{4,d_0,d_1}(t, c_0, c_1)$. Then by (25), one of the following cases obtains:

$$c'_1 \neq c_1; \tag{39}$$

$$\rho(t', c_0) > \max(d_0, \rho(t, c_0)); \tag{40}$$

$$\rho(c'_0, c_1) > \max(d_1, \rho(c_0, c_1)). \tag{41}$$

In case (39), c'_1 is not in $\{c_1\}$. In case (40), t' is not in $T_{4,d_0,d_1}(t, c_0, c_1)$. In case (41), c'_0 is not in $C_{4,d_0,d_1}(t, c_0, c_1)$. $\square$

Worlds are three dimensional. The preceding proposition says that $R_{4,d_0,d_1}(t, c_0, c_1)$ is a two-dimensional, finite rectangle whose sides are the known intervals for t and for c_0. Furthermore, it is easy to verify from definition (53) that:

Proposition 4.

$$|T_{4,d_0,d_1}(t, c_0, c_1)| = 2\max(d_0, \rho(t, c_0)) + 1; \tag{42}$$

$$|C_{4,d_0,d_1}(t, c_0, c_1)| = 2\max(d_1, \rho(c_0, c_1)) + 1. \tag{43}$$

Now one may calculate:

$$|R_{4,d_0,d_1}(t, c_0, c_1)| = |T_{4,d_0,d_1}(t, c_0, c_1) \times C_{4,d_0,0}(t, c_0, c_1) \times \{c_1\}| \tag{44}$$

$$= |T_{4,d_0,d_1}(t, c_0, c_1)| \cdot |C_{4,d_0,0}(t, c_0, c_1)|\,. \tag{45}$$

The important point for Williamson is that, in our model, the centering parameter c_0 performs the same function that t served in Williamson's first model—no value of c_0 other than the actual value allows for knowledge that one knows when S's confidence interval is true (i.e., in the very probable case in which $\rho(t, c_0) \leq d_0$).

Proposition 5. *Suppose that $\rho(t, c_0) \leq d_0$. Then: $R_{4,d_0,d_1}(t', c_0', c_1') \subseteq R_{4,d_0,0}(t, c_0, c_1) \Rightarrow c_0 = c_0'$.*

Proof. Immediate consequence of proposition 3 and (12). □

The preceding proposition is false when $\rho(t, c_0) \leq d_0 > d_0$.

Therefore, in light of propositions 5 3 and (12), we have the following result when $\rho(t, c_0) \leq d_0 > d_0$:

$$|KR_{4,d_0,d_1}(t, c_0, c_1)| \leq |T_{4,d_0,d_1}(t, c_0, c_1) \times \{c_0\} \times \{c_1\}| \tag{46}$$

$$= |T_{4,d_0,d_1}(t, c_0, c_1)|. \tag{47}$$

Hence, the evidential probability of $KR_{4,d_0,d_1}(t, c_0, c_1)$ when $\rho(t, c_0) \leq d_0 > d_0$ is no greater than:

$$\frac{T_{4,d_0,d_1}(t, c_0, c_1)}{T_{4,d_0,d_1}(t, c_0, c_1) C_{4,d_0,0}(t, c_0, c_1)} = \frac{1}{|C_{4,d_0,0}(t, c_0, c_1)|} \tag{48}$$

$$= \frac{1}{2\max(d_1, \rho(c_0, c_1)) + 1}. \tag{49}$$

The smallest possible concession the luminist can make to Williamson short of full validation of the KK principle is to set $d_1 = 1$. But then the preceding expression for evidential probability is at most 1/3 and the failure of KK becomes even larger as d_1 increases. So by Williamson's criterion, the failure of the KK principle is very likely to be large, even given the minimum uncertainty about the position of of one's confidence interval over clock readings. The minimum uncertainty can be made arbitrarily small, in terms of angular separation, by increasing the number of potential clock readings without increasing the size of the clock. That is a surprising and apparently uncomfortable result for moderate luminists.

13 Soundness of the Sorites Argument

ETC

14 Critique of Centering Ignorance

It is difficult to interpret S's ignorance of the center of her own confidence interval in the preceding model as anything but an assumed aphasia in S. Perhaps the interval is latent and dispositional, whereas S's knowledge thereof is explicitly propositional, and explicit cognition has trouble eliciting latent cognition. But then why shouldn't dispositional knowledge of latent, dispositional knowledge should suffice for KK? Moreover, it is risky to conclude on the basis of armchair intuitions that we don't have latent, second-order knowledge on the basis of manifest intuitions—some serious, cognitive neuroscience is required. Furthermore, if a scientist who publishes an explicit confidence interval believes what she publishes and performed the statistical rituals properly, it seems that she has access to her known interval, so one might say that anyone with access to a statistics text is "in a position" to avoid the failure of KK in the preceding clock model. In any event, it should hardly come as a surprise that S fails to know that she knows if she is assumed not even to believe that she knows what she knows. For that reason, Williamson explicitly disavows arguments for non-luminosity that are based on an explicit denial that S believes that she believes, so apparently he would not be much comforted by the non-luminosity result obtained in that model.

It remains impressive that the smallest possible failure to believe that one believes results in a large failure to know that one knows in our model. But that result is also less interesting than it might seem. Williamson's original examples are shocking because the alleged non-luminosity they illustrate concerns ordinary perceptual knowledge about the clock reading t. In the preceding model, the non-luminosity concerns the entire knowledge state $K_{4,d_0,d_1}(t, c_0, c_1)$, which involves the relative position of c_0 and c_1 as well as the position of t. If we focus, as Williamson does, on the perceived interval $T_{4,d_0,d_1}(t, c_0, c_1)$ around t, then the probability that S knows that she knows her perceived interval is just:

$$
\begin{aligned}
p(KT_{4,d_0,d_1}(t,c_0,c_1) \mid T_{4,d_0,d_1}(t,c_0,c_1)) &= \frac{|KT_{4,d_0,d_1}(t,c_0,c_1)|}{|T_{4,d_0,d_1}(t,c_0,c_1)|} && (50)\\
&= |T_{4,d_0,d_1}(t,c_0,c_1)| - 2d_0\, |T_{4,d_0,d_1}(t,c_0,c_1)| && (51)\\
&= \frac{2d_2+2d_1+1}{2d_1+1}. && (52)
\end{aligned}
$$

If, plausibly, the aphasic quantity $2d_2$ is small compared with the confidence interval width $2d_0+1$, then the degree of failure of KK with respect to position is also small, so even if Williamson were to embrace aphasia as an argument for non-luminosity, the amount of non-luminosity obtained is pretty much just the amount of assumed aphasia—hardly a just cause for alarm.

15 Empirical Non-Luminosity

A more plausible and uncontroversial source of non-luminosity concerns the *width* of S's confidence interval. In the case of observational noise, the width of the interval depends on the variance or spread of the approximately Gaussian distribution over observational errors, and that parameter must be estimated empirically, for the situation at hand. That empirical estimate is also noisy, resulting in a meta-confidence-interval over d_0. Let c_2 be the center of that interval and let its width be $2d_2+1$. Then, it is natural to define accessibility as follows:

$$
R_{5,d_0,d_2}((t,c_0,c_2),(t',c_0',c_2')) \iff \begin{cases} c_1' &= c_1; \\ \rho(t',c_0) &\leq \max(d_0,\rho(t,c_0)); \\ \rho(d_0',c_2) &\leq \max(d_2,\rho(d_0,d_2)). \end{cases} \quad (53)
$$