Stalnaker on the KK principle

Andrew Bacon*

September 4, 2013

Abstract

In a recent exchange with Hawthorne and Magidor, Stalnaker has provided an interesting objection to Williamson's margin of error principle that could form the basis of a defense of the KK thesis. I argue that Stalnaker's position is more robust than Hawthorne and Magidor's recent reply suggests. I instead present an alternative refutation of the KK principle which does not appeal to the margin of error principle.

In a recent exchange with Hawthorne and Magidor, Stalnaker [4] has provided an interesting defense of the KK thesis[1]:

KK If one knows that p, then one is always in a position to know that one knows that p.

The KK principle is now widely thought to face difficulties when ones knowledge of p's subject matter is 'inexact' (see Williamson [5], [6].) A paradigm example of inexact knowledge is knowledge gained from an imprecise perceptual experience. Following [6], suppose that Mr. Magoo sees a tree from a distance, and thereby gains inexact knowledge concerning the height of the tree. He certainly gains some knowledge from the experience – for example, he now knows the height of the tree is less than 1000cm, and he knows it's more that 10cm. However he does not know whether the tree is larger or smaller than 200cm. Following Williamson, we may also suppose a margin for error principle is true of Mr. Magoo, and that he knows it. I.e.:

ME Mr. Magoo knows that if the tree is $n$cm tall, he does not know that it is less than $n + 1$cm tall.

On the basis of ME, the KK principle, some natural closure principles and the fact that Mr. Magoo knows the tree is larger that 10cm and smaller than 1000cm one can derive a contradiction.[2] Stalnaker's target is Magoo's knowledge of the margin for error principle. He imagines that Mr. Magoo is cross examined about

*Acknowledgements

[1]The discussion was primarily concerned with the propositional attitude 'presupposes that', however, both parties discuss and acknowledge the importance of these considerations for the KK principle.

[2]See [6].

his epistemic situation, and finds the following dialogue between Mr. Magoo and his interviewer reasonable[3]

I: How tall do you think the tree is.

Mr. M: If forced to put a number on it, I would guess the tree to be 200cm.

I: do you think the tree is exactly 200cm.

Mr. M: No, but I know that the tree is between 100cm and 300cm. In general, I know that my best guess is reliably within 100cm of the actual height of the tree. So I know that when I guess the tree to be 200cm then the tree is between 100cm and 300cm.

I: Might the tree be 300cm.

Mr. M: For all I know, the tree is 300cm. Although I think this is highly improbable since it is so far from my estimate.

Assuming that Mr. Magoo's responses were knowledgeable, we may infer the following things

P1. That Mr. Magoo knows that his best guess is reliably, and thus knowledgably, accurate to within 100cm.[4]

P2. Mr. Magoo knows that his best guess is 200cm.

P3. For all Mr. Magoo knows the tree is 300cm.

From this and closure principles we may infer:

> It is compatible with Mr. Magoo's knowledge (but highly unlikely) that the following is true: his best guess is 200cm, but the tree is 300cm.

If we assume also that Mr. Magoo knows P2 we may infer

¬ME It is compatible with Mr. Magoo's knowledge that: the tree is 300cm and he knows that it's less that 301cm.

However, this directly contradicts the claim that Mr. Magoo knows the margin of error principle.

Note that Stalnaker's argument does not refute the margin of error principle itself. Mr. Magoo may in fact satisfy it, but fail to know it. Note also that according to Stalnaker's description of the case Mr. Magoo finds each instance of the margin of error principle highly probable. For example Stalnaker writes: "the situation [that my best guess is 200cm, but the tree is 300cm] is extremely improbable, since if the actual height were that far above 200cm, it is almost

[3] I have adapted this from Stalnaker's dialogue [4].

[4] In other words, he knows that if he guesses the tree to be $n$cm, the tree is between $n - 100$cm and $n + 100$cm, and he knows this fact about himself.

certain that my best guess would be higher.[5] So the situation is highly unlikely, but I agree that I can't rule this possibility out absolutely."[6] Thus, according to Stalnaker, the failures of the margin of error principle are all highly improbable, although they cannot be ruled out absolutely by Mr. Magoo's knowledge.

1 Discussion

Let us now turn to the principles P1-P3. In order to get ¬ME, on my reconstruction of the argument, one needs to assume additionally that Mr. Magoo knows P2, and one needs to assume some closure principles. To give the argument its best shot I take it for granted that if the claim that Mr. Magoo knows his best guess is 200cm is defensible, so is the claim that he knows he knows this. However, this assumption is needed: it could be that Magoo in fact knows his best guess and that the tree is within 100cm of that guess, but in the epistemically accessible worlds where the tree is 300cm he does not know this is his best guess, and thus does not know that the tree is less that 301cm.

P2 might seem innocuous, but its tenability is highly dependent on what one means by 'best guess.' For example, suppose that whenever Mr. Magoo sees a tree that is n meters from a distance, the probability distribution on his evidence over possible heights forms a smooth curve whose peak is at n, and which trails off to 0 at $n \pm 100$cm. Then P1 is true if 'best guess' is taken to mean the peak of Magoo's probability curve. However, it is far from clear, and slightly question begging in this context, to assume that his best guess is introspectible if we understand 'best guess' in this way.

I think it is clear that this is not what Stalnaker means by 'best guess', it is rather the number that Mr. Magoo would say if forced to say a precise number. It might be that there is no such number, or it is indeterminate which number he would say, if there were several equally close worlds where he answers differently. In that case let us say that his best guess is the number he is most likely to say if asked.

One might rightly question whether even this is introspectible. What if, for example, he is disposed to say 200cm 51% of the time and 201cm 49% of the time. It seems like it would be very hard for Mr. Magoo to know he was more

[5] I take it that in Stalnaker's picture, Mr. Magoo's probability distribution over possible tree heights after his visual experience is smooth curve with a peak at around 200cm, which approaches zero at around 100cm either side; thus $Pr(h = 300) \approx 0$. This is in contrast to Williamson's analysis ([REF] (very?) improbable knowing) where the probability curve is a sharp rectangle (assuming uniform priors) obtained by conditioning on Magoo's knowledge. One might reconcile the two pictures by including further epistemic states in the model: perhaps the strongest proposition Mr. Magoo learns after seeing the tree is a *vague* proposition. Perhaps he sees that the tree is about 200cm and conditions on this proposition. Since it is vague whether the tree can be both 210cm and about 200cm, it is unknown either way, thus there must be two epistemic states: one where trees that are 210cm are about 200cm and one where they're not. This would naturally give rise to a situation where conditioning uniform priors on your knowledge gives a smooth curve, without invoking Jeffrey conditioning.

[6] I have changed numbers to fit the current example.

disposed to say 200cm than 201cm.[7] But even granting that he knows which he is more likely to say, why should P1 be true on this reading of 'best guess.' If Magoo's probability distribution really is a curve whose support is the range $n \pm 100$cm, where n is the peak of his distribution, then why should he know the tree is between $g - 100$ and $g + 100$, where g is the number he would say if asked? g is obtained from n, the peak of his probability curve, by rounding it up or down in some psychologically plausible way (I take it anyone realistic about their own discriminatory powers will only answer in whole numbers), and this rounding operation may shift the range left or right in such a way that it overlaps with epistemically inaccessible worlds.

How essential is P2 to the argument? Recall that all one needs to refute ME is (i) that Mr. Magoo knows he knows the tree is between 100cm and 300cm, and (ii) that it's compatible with his knowledge that the tree is 300cm. Invoking Mr. Magoo's knowledge about the accuracy of his best guess was a way of making (i) sound plausible – but isn't (i) itself antecedently plausible given Mr. Magoo's responses? In my view it would not greatly damage the plausibility of the argument if we took (i) as a premise instead. There are, however, more pressing objections so I shall leave this issue for the time being.

Let us turn now to P1 and P3. In [3] Hawthorne and Magidor argue that in most realistic situations either P1 or P3 will fail. Let us suppose that P1 holds, so that Magoo knows that his best guess is reliably within 100cm of the actual height of the tree. In other words, he knows that whenever his best guess is 200cm, the tree is between 100cm and 300cm. They argue that in most realistic situations Mr. Magoo's guess will actually be reliably accurate to within a smaller margin than this, although he may not know this. In other words, he will not know the strongest true proposition of the form: whenever Mr. Magoo's best guess is $n$cm, the tree is between $n - x$cm and $n + x$cm. Let us suppose the strongest true propositon is obtained when $x = 98.76$. Although he may well know the weaker proposition that his best guess is knowledgeably within 100cm of the trees actual height, he is *in fact* reliable to within 98.76cm. Since his best guess is 200cm we thus have:

> Mr. Magoo knows the tree is between 101.24cm and 298.76cm.

Now, since this knowledge actually rules out the possibility that the tree is 300cm (or even 298.77cm), it is not compatible with his knowledge that the tree is 300cm, contradicting P3. On the other hand, it *is* compatible with his knowledge that the tree is 298.76cm. Could we not substitute 298.76cm for 300cm in P3 and modify our argument so that P1 reads: "Mr. Magoo knows that his best guess is reliably, and thus knowledgably, accurate to within 98.76cm"? While P3 is true in this modification, H&M claims that this version of P1 is not; to know that his guess is knowledgably accurate to within 98.76cm requires Mr. Magoo to have implausibly precise knowledge about his own discriminatory powers, they claim.

[7] It would, for example, take hundreds of experiments to show with high probability that there's a higher chance of him saying 200cm.

Let us consider this argument in more detail. We may simplify H&M's response by noting that, since Mr. Magoo's guess is reliable within 98.76cm but not reliable at shorter distances, 1-4 are all true, but Mr. Magoo plausibly only knows 1, and fails to know 2, 3 or 4.

1. Mr. Magoo knows that the tree is between 100cm and 300cm.
2. Mr. Magoo knows that the tree is between 101.24cm and 298.76cm.
3. Mr. Magoo does not know that the tree is between 101.25cm and 298.75cm.
4. 2. (and not 1.) is the strongest true proposition of the form: Mr. Magoo knows that the tree is between 200-$x$cm and 200+$x$cm.

H&M's defence of ME relies on knowledge about the actual margin of error from ones best guess being hard to come by, and therefore that knowledge of 2 is hard to come by. However, without further reason to think that Mr. Magoo does not know 2, the dice seem to be loaded heavily against Stalnaker. For if Mr. Magoo does not know 2 then the KK principle is already refuted, and the need for a margin of error principle is circumvented altogether. To bolster this intuition they contrast the case to an unrealistic case in which Mr. Magoo is a scientist who has performed extensive experiments and has been able to determine that his best guess is always at most 98.76cm from the trees actual height. This kind of knowledge certainly *does* seem hard to come by; determining x to two decimal places would require many experiments, or some kind of extraordinary ability. Even with experiments it seems unlikely that Mr. Magoo could know his margin of error to that degree of accuracy.

However, I think it is clear that Mr. Magoo's epistemic situation is not analogous to that of the scientist. For, unlike the scientist, Mr. Magoo does not know 3. or 4. If asked how accurate he thought his guesses are, he might, if pushed, assent to the figure of 98.76cm (after all, there will always be *some* point at which he stops assenting) but be unsure if that's his margin of error because he's unsure whether he also knows that the tree is between 101.25cm and 298.75cm. Unlike the scientist, Mr. Magoo is not able to place absurdly precise bounds on his powers of discrimination. He does not know which number in the interval $[90, 98.76]$ his margin of error is, so unlike the scientist he does not even know to the nearest integer what his margin of error is. Mr. Magoo may know that his guess is reliable to within 98.76cm, without invoking extraordinary knowledge about his own discriminatory abilities.[8] It is rather that one needs

[8]If one is thrown by the apparent precision of the number 98.76, note that he knows many very precise things about his powers of discrimination: for example he knows that the tree is always 11290.519098cm from his best guess. It is knowledge that this is the strongest bound on his discriminatory powers that is hard to come by. (There is an issue which I have for the most part been ignoring: whether it is psychologically plausible that Mr. Magoo even *believes* exceedingly precise propositions, like the proposition that the tree is always 11290.519098cm from his best guess, even if they follow from his other beliefs. In discussions of the KK principle in the context of epistemic logic this issue is often set aside; I shall continue to set aside this issue for now. It may be hard to know certain things because it is difficult, psychologically speaking, to form specific precise beliefs, but it is not this kind of difficulty that is playing an important role in H&M's argument.)

the extraordinary abilities to know, in addition, that this is the smallest distance within which his guesses are reliable.

This would all be impossible if we subscribed to the negative introspection principle for knowledge:

K¬K If one does not know that p, then one is always in a position to know that one does not know that p

Both Hawthorne and Magidor and Stalnaker agree that negative introspection is implausible. Without negative introspection, however, it is completely consistent to say that Mr. Magoo knows that his guess is reliable to within 98.76cm, even though he would proclaim ignorance if we asked him to specify the smallest such number.

2 Knowing your discriminatory powers

If we are to draw anything from the previous discussion, it is that there seems to be a gap in the anti-KK arguments based on the margin for error principle. Stalnaker has outlined a natural looking view in which the margin for error principle, while highly probably, is not known, and in which the KK thesis appears to be consistent, H&M's argument notwithstanding. In the rest of the paper I shall outline another argument against the KK principle. This argument is premised on H&M's thesis that in ordinary cases people do not know the exact extent of their discriminatory powers; the argument does not appeal to ME or the negative introspection principle.

Our premise thus entitles us to assume, with Hawthorne and Magidor, that the epistemic state of H&M's scientist is rare and hard to come by, and we believe that Mr. Magoo in this case is No Scientist[9]:

NS For no i, does Mr. Magoo know that he knows the tree is less than $i+1$cm and know that he doesn't know that the tree is less than $i$cm.

If we assumed the KK thesis and negative introspection this principle would be inconsistent. But we are not interested in defending negative introspection. Furthermore, the case for this principle seems quite strong: it would require extraordinary circumstances for me to be able to detect the exact number, i, at which I stop knowing that the tree is less than $i$cm.[10] I would need perfect knowledge of my discriminatory powers.

[9] Analogies between this principle and 'gap principles' from the literature on higher order vagueness can be drawn here. Arguments analogous to the ones I make here are run against variations of the principle below in [7], [1] and [8]. A couple of salient differences are worth remarking upon, however: some principles from that literature are clearly not motivated when one is reading NS, as I am, as a principle about what a particular agent knows (an analogue of the principle which allows one to infer that Mr. Magoo knows that p, from p is assumed in [7] and [1], for example.) On the other hand those arguments are not concerned with the analogue of the KK principle; a fact which allows us to circumvent the use of the unmotivated principles in this particular context.

[10] Note of course that this argument can be run in different units; thus if one found this not totally implausible, consider the analogous case for picometres.

On the face of it, this principle is also weaker than the margin for error principle. Since these considerations seem to be perfectly general it is natural to assume that Mr. Magoo knows each instance of NS. That is, for each i, Mr. Magoo knows that he doesn't both know he knows the tree is less than $i+1$cm and know he doesn't know the tree is less than $i$cm.

Let p_i be the proposition that the height of the tree is less than or equal to $i$cm. Let us write Kp to mean 'Mr. Magoo knows that p' and Mp to mean 'it is compatible with Mr. Magoo's knowledge that p.' I shall assume these are interdefinable so that $Mp := \neg K \neg p$. Our principle can be encoded by the schema: $\neg(KKp_{i+1} \wedge K\neg Kp_i)$, or equivalently $KKp_{i+1} \rightarrow MKp_i$. The claim that Mr. Magoo knows each instance of this schema can be expressed with another schema: $K(KKp_{i+1} \rightarrow MKp_i)$. It is very natural to see if this principle is compatible with $\mathsf{S4}$:

CL $\phi[\psi/p]$ whenever ϕ is a tautology of propositional logic.

K $K(\phi \rightarrow \psi) \rightarrow (K\phi \rightarrow K\psi)$.

T $K\phi \rightarrow \phi$.

KK $K\phi \rightarrow KK\phi$

Nec. If ϕ is a theorem of $\mathsf{S4}$ so is $K\phi$.

MP. If ϕ and $\phi \rightarrow \psi$ are theorems of $\mathsf{S4}$ so is ψ.

Unfortunately, it is not too hard to see that it is not consistent in $\mathsf{S4}$ to hold that Mr. Magoo knows each instance of NS, knows that the tree is greater than 0cm and not greater than 1000cm.

Theorem 2.1. *The following is inconsistent in* $\mathsf{S4}$

$K(KKp_{i+1} \rightarrow MKp_i)$ *for each* $i \leq 1000$

Kp_{1000}

$K\neg p_0$

Let me end by flagging a limitation of this theorem. The system KT, where only K and T are axioms, with the KK principle as a true non-axiom is strictly weaker than the logic $\mathsf{S4}$. $\mathsf{S4}$ allows one to infer that, as well as knowing each theorem of KT, one knows each theorem of $\mathsf{S4}$, including each instance of the KK principle. This is important since a proponent of the KK principle is not committed to this aspect of full $\mathsf{S4}$. The KK thesis represents a way of evaluating an agents propositional attitudes as good (rational) or bad (not rational) according to whether they satisfy certain conditions, such as knowing something only if they know they know it. The range of applicability of the KK principle is agents who are epistemically good in this sense, and is not supposed to be a principle governing what the average irrational person knows and believes.

However, whether the circumstances are good for an agent need not be accessible to the agent. The good agents, those who are doing well epistemically, may know that all good agents satisfy the KK principle, but not know that they themselves are good. In such a case it would not be reasonable to assume that one must always know that one satisfies the KK thesis.

3 Appendix

In this appendix theorem 2.1 is shown. $K^n p$ is short for n 'K's followed by a 'p'; similarly for $M^n p$. We begin by listing two standard facts about S4.

Lemma 3.1. *$K\phi$ and $K^n\phi$ are intersubstitutable in all contexts.*

Thus in particular since we have $K(KKp_{i+1} \to MKp_i)$ we have every instance of (i) $K(KK^n p_{i+1} \to MK^n p_i)$

Lemma 3.2. $K(\phi \to \psi) \to (M^n\phi \to M^n\psi)$.

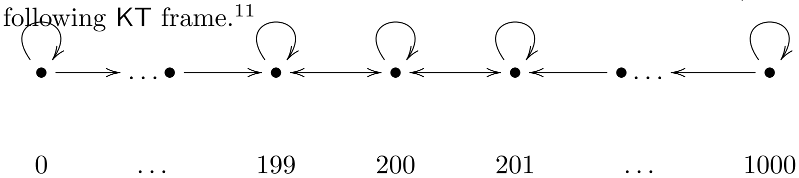
Since we have (i) $K(KK^n p_{i+1} \to MK^n p_i)$, we also have every instance of (ii) $M^n KK^m p_{i+1} \to M^n MK^m p_i$. Our argument now proceeds:

1. $Kp_{1000} \to K^{1000}p_{1000}$ (by the KK thesis $\times$ 1000)
2. $KK^{999}p_{1000} \to MK^{999}p_{999}$ (by (ii))
3. $MKK^{998}p_{999} \to MMK^{998}p_{998}$ (by (ii))
4. $MMKK^{997}p_{998} \to MMMK^{997}p_{997}$ (by (ii))
5. ...
6. $M^{999}KKp_1 \to M^{1000}Kp_0$
7. So by transitivity of $\to$ we have $Kp_{1000} \to M^{1000}Kp_0$
8. $M^{1000}Kp_0$ by modus ponens

Now note that $K(K\neg p_0 \to \neg Kp_0)$, so by K we have $K\neg Kp_0$ and by the KK thesis: $K^{1000}\neg Kp_0$. This is inconsistent with $M^{1000}Kp_0$ by the definition of M.

Theorem 3.3. *$K(KKp_{i+1} \to MKp_i)$ for each $i < 1000$, Kp_{1000}, $K\neg p_0$ and $K\phi \to KK\phi$ are consistent in* KT

Note that all the principles listed are true at the middle point (i.e. 200) of the following KT frame.[11]



[11] For the diagram's sake I have reduced Mr. Magoo's margin of error from 100 to 1. It is marginally more complicated to modify this to a more realistic margin of error.

It's worth pointing out that the full power of $\mathsf{S4}$ is only needed for proving (i). If we assumed (i) instead of $K(KKp_{i+1} \rightarrow MKp_i)$, we could arrive back at a contradiction in KT with KK as a premise.

References

[1] D.G. Fara. Gap principles, penumbral consequence, and infinitely higher-order vagueness. *Liars and Heaps: New Essays on Paradox*, pages 195–222, 2003.

[2] J. Hawthorne and O. Magidor. Assertion, context, and epistemic accessibility. *Mind*, 118(470):377, 2009.

[3] J. Hawthorne and O. Magidor. Assertion and epistemic opacity. *Mind*, 119(476):1087, 2010.

[4] R.C. Stalnaker. On hawthorne and magidor on assertion, context, and epistemic accessibility. *Mind*, 118(470):399, 2009.

[5] T. Williamson. Inexact knowledge. *Mind*, 101(402):217–242, 1992.

[6] T. Williamson. *Knowledge and its Limits*. Oxford University Press, 2002.

[7] C. Wright. Is higher order vagueness coherent? *Analysis*, 52(3):129, 1992.

[8] E. Zardini. Squeezing and stretching: How vagueness can outrun borderlineness. In *Proceedings of the Aristotelian Society (Hardback)*, volume 106, pages 421–428. John Wiley & Sons, 2006.