

A Problem about Permission and Possibility¹

Stephen Yablo

¹ I am grateful to Seth Yalcin, Eric Swanson, Bob Stalnaker, Peter Railton, Allan Gibbard, Rich Thomason, Thony Gillies, Kai von Fintel, Sarah Moss, Sally Haslanger, Caspar Hare, Carrie Jenkins, Kit Fine, George Bealer, Andy Egan, Brian Weatherson, John MacFarlane, Tim Williamson, Gideon Rosen, Bob Hale, Elizabeth Fricker, Frank Jackson, Crispin Wright, Josh Dever, Nicholas Asher, Adam Morton, David Beaver, Cleo Condoravdi, Danny Fox, Brian Hedden, Nina Emery, Daniel Greco, Susanna Rinard, Andrew Graham, and two anonymous referees for this volume. Bob Stalnaker was kind enough to show me remarks he made in response to Lewis's presentation of his permission problem at Cornell. Some of this material figured in my 2007 Nelson Lectures at Michigan and 2008 Hempel Lectures at Princeton; thanks to the audiences at both places. I learned a lot from the participants, beyond those already mentioned, in the 2006 Austin Workshop on Epistemic Modals and the 2006 Arché Conference on Modality; particular thanks to my Arché commentator David Efird. This paper descends from a paper appearing, with David's comments, in an Oxford volume on modality edited by Bob Hale.

This paper explores the prospects for a unified theory of deontic and (so-called) epistemic modality. That we use similar language for the two sorts of modality is a familiar point. The audiovisual system made the point for me the last time I gave this paper; while warming up it projected these words onto the screen:

If this equipment is off campus, it may be stolen.

The intended reading, of course, is that a certain hypothesis is not ruled out: that the projector is stolen. But another possible reading is that a certain course of action is not ruled out: that the projector be stolen.² Compare in this respect "Sabotage is not ruled out," as uttered by an FAA investigator after the crash, to the same sentence uttered by rebel leader Natasha before the crash. The investigator is saying that sabotage is not ruled out descriptively, as it would be if he'd asserted "There was no sabotage." Natasha is saying that it is not ruled out prescriptively, as it would be if she'd commanded her underlings not to engage in sabotage. That the two readings of "may" correspond to two readings of "not ruled out" is suggestive; it suggests that "That may be so" and "That may be done" have semantic properties in common.

² Of course, it is a different sort of ruling out that's intended; stealing the projector is not forbidden.

What does it matter, though, if the one "may" has properties in common with the other? It matters because descriptive "may" is extremely confusing; and the questions we are driven to as we attempt to understand it are questions that, as it happens, have been much discussed in connection with deontic "may."

The standard semantics for "It may (or might, or is possible) that φ "³ has it expressing something in the vicinity of the speaker's failing to know that \sim . Thus Moore:

It's possible that I'm not sitting down now....means 'It is not certain that I am' or 'I don't know that I am'⁴

Later versions of the standard semantics allow the knower(s) and/or the information against which φ is tested to vary:⁵

It is possible _{Δ} that p is true if and only if what Δ knows does not, in a manner that is obvious to Δ , entail not- p .⁶

³ I will generally use "might" rather than "may" when the (so-called) epistemic reading is intended.

⁴ Moore 1962, p184.

⁵ DeRose 1991; Hacking 1967; Teller 1972; von Fintel and Gillies 2007.

"It is possible that p " is true if and only if (1) no member of the relevant community know that p is false, and (2) there is no relevant way by which members of the relevant community can come to know that p is false.⁷

There is undoubtedly something right about this approach. But there are things wrong with it too.

One problem with the standard semantics is that it gets the subject matter wrong. When I say, "Bob might be in his office," I am talking about Bob and his office, not myself or the extent of my information.⁸ The difference in

⁶ Stanley 2005, p128.

⁷ DeRose 1991, p593-4.

⁸ It might seem the worry can be sidestepped by putting speaker's knowledge into the mechanism by which the content is generated, rather than the content itself (Kratzer 1981). Speaker's knowledge would play the same sort of role in the evaluation of "might"-claims as speaker's attention plays in the evaluation of "you"-claims. Andy Egan pointed out that this solution may swing too far in the other direction. For the Kratzer-proposition can be seen as a consistency-claim: φ is consistent with K , where K is the relevant body of knowledge. The modal and evidential properties of "might"-claims look rather different from those of consistency-claims. Clouds are evidence that it might rain, but not evidence

subject matter makes for a difference in attitude. Imagine that the building is on fire and everyone other than Bob has escaped down the back stairs. I am afraid that Bob might still be in his office. I am not afraid that I don't know he's elsewhere.

Two, the proposed truth-conditions in their naïve Moorean form are too weak. The mere fact that I don't myself know that $\sim\varphi$ doesn't make it true in my mouth that φ might be so. Suppose you question my claim on the basis that Bob was just seen stepping onto a plane. It would be no reply at all to say that my information really was as limited as I suggested; I really and truly didn't know that Bob was not in his office. Evidently the information that needs to comport with φ for a might-claim to be correct can extend beyond what the speaker personally knows at the time of utterance. The proposal should really be that "It might be that φ " is true iff φ is not ruled out by any pertinent facts – where the test of pertinence is, presumably, that the speaker is prepared to acknowledge that she was mistaken if these facts really do/did obtain.⁹

for the consistency of φ with K. Consistency-claims are necessary or impossible, but the gathering clouds hardly make it a necessary truth that it might rain.

⁹ For more on these issues, see Egan 2007; Egan, Hawthorne, and Weatherson 2005; MacFarlane 2006; and von Fintel and Gillies 2007.

But, and this is the third problem, the truth-conditions are now so strong that speakers generally have no idea whether they are satisfied; they have no business, then, asserting that φ might be so. The principle here is that I should not assert that χ , if (a) I am aware of a ψ such that χ is false if ψ is true, and (b) I consider ψ entirely likely to be true.¹⁰ When χ is "It might be that φ ," I am virtually always aware of a ψ like that, viz. "Somewhere out there, there is evidence that rules φ out." ψ meets condition (a) because I freely accept that my might-claim is mistaken if φ is ruled out by the evidence, including evidence I don't myself possess. (I freely accept, for instance, that if Bob was seen getting on a plane at 11:55, then it is not true after all that he might now (at noon) be in his office.) ψ meets condition (b) because I do not, when I say that it might be that φ , take myself to know that my evidence is relevantly complete; obviously I might be missing something which makes φ unlikely.¹¹ (When I say that Bob might be in his office, I do not take myself to know that no one has just seen him get on a plane.) The problem is this: if I think it entirely likely that there is evidence that exposes my statement as false, how in good conscience can I make the statement? Who would dare make a might-claim, if the claim was entirely likely to be mistaken?

¹⁰ Sarah Moss uses a principle like this in her work on reverse Sobel sequences.

¹¹ If that kind of knowledge were required, I would not say "It might be that φ and it might be that $\sim\varphi$ " unless I thought that φ was objectively undecidable, in the sense that all the evidence in the world left it an open question whether φ .

The fourth problem with the standard semantics is that it is too epistemic. I have a thing about the sanctity of the ballot box, imagine, so when you ask me whether I am going to vote for Kucinich, I say, "I might or I might not," despite knowing perfectly well what I've decided, and not trying to hide the fact that I know.¹² Or suppose that I run into a creditor who demands that I give him a check by the end of the day. I know perfectly well that I am going to do what he asks – I have the check in my pocket – but still I say, "I might or I might not; it might have to wait until tomorrow," for the loan is not strictly due until Friday. I say it not because I don't know I'll give him the check today, but because I reserve the right not to; it's not a limit in my information I'm indicating, but a limit in what I'm prepared to commit to. One final example. Imagine I am pitching a story line to a Hollywood mogul. "Now comes the good part," I tell him. "The Raskolnikov character brutally murders the pawnbroker." "Not a chance, not if we want PG-13," comes the reply. "OK," I say in a concessive spirit, "so he might just rough her up a bit." The "might" here is not to mark the limits of my knowledge. There is nothing to know in this case; we are tossing around ideas. I say "might just rough her up" to convey that I am scaling back the proposal.

A fifth problem is more logical in nature. Suppose that φ is consistent with all pertinent information. Then so is everything φ entails. One would expect, then, that if φ entailed ψ , "It might be that φ " would entail "It might

¹² If I say I don't know whether I'll vote for Kucinich, you can properly accuse me of lying. But it is not a lie to say I might vote for him or I might not.

be that ψ ." "Bob might be in his office," for instance, should entail "Bob might be in his office or in an opium den." And yet "Bob might be in his office or in an opium den" seems to make a stronger claim, roughly to the effect that Bob might be at the one place and in addition he might be at the other. There is of course a similar puzzle about permission: how is that "You can go or stay" entails (or seems to) that you can do whichever you want, that is, it is open to you to go and it is open to you to stay.¹³

The sixth problem I learned from MIT graduate student¹⁴ Seth Yalcin. An advantage sometimes claimed for the standard semantics is that it explains the paradoxicality of " φ and it might be that $\sim\varphi$." To say that φ and it might be that $\sim\varphi$ is to say this: φ but I don't know that φ . And the problem with " φ but I don't know that φ " we more or less understand; it's an instance of Moore's paradox. But there are reasons to doubt that " φ and it might be that $\sim\varphi$ " is no more than a form of Moore's paradox. That paradox is thought to arise because the Moore-sentence is unassertable. The proof that it's a problem of assertability rather than truth is that there is nothing to stop me from supposing, in the antecedent of a conditional, that φ and I don't know it: If I am unbeknownst to myself dreaming that I am typing this paper, then I am most likely at home in bed. If " φ and it might be that $\sim\varphi$ " were paradoxical only for Moorean reasons, one would expect it to be supposable too. And it isn't. "If I am dreaming that I am typing but might not be dreaming that I am typing ..." makes no sense. The sixth problem is

¹³ Kamp 1973; Zimmermann 2000; Fox 2006.

¹⁴ At the time of writing—now teaching at NYU.

that " φ and it might be that $\sim\varphi$ " is not coherently supposable, and the standard semantics offers no explanation of this.¹⁵

The traditional "static" semantics for epistemic modals has its problems. This has led some to propose a dynamic semantics; the meaning of "might φ " is given not by its truth-conditions but its effect on context or shared information. The best-known version of this is Frank Veltman's default semantics.¹⁶ Veltman thinks of "might φ " as checking on φ 's consistency with the going information state. Uttered in information state S , it returns S if S is consistent with φ , and it returns the null information state if S is not consistent with φ . Both parts of this seem prima facie at odds with the way "might" is used.

¹⁵ Yalcin thinks that there is a problem here for any truth-conditional semantics. For the following is an invalid argument: *It might be the case that $\sim A$, therefore $\sim A$* . A one-premise argument *X, therefore Y* is invalid, one would think, only if the conclusion can be false while the premise is true--only if there is a possible scenario where $\neg Y \& X$. In this case, that means a possible scenario where *A and it might be that $\sim A$* . And there is no such scenario.

¹⁶ Gillies 2004; Veltman 1996.

Consider first the idea that "might φ " returns the null state when S is inconsistent with φ . Suppose it is understood all around that John, Paul, George, and Ringo will be at the party. Then someone runs in with the news that Ringo might not be able to make it. Ringo's not making it is inconsistent with John, Paul, George, and Ringo being there. All our information is demolished, then, according to default semantics. But the information that John, Paul, and George will be at the party surely remains when we learn that Ringo might not attend.¹⁷

The idea that "might φ " returns S if φ is consistent with S seems questionable, too. It's consistent with John, Paul, George, and Ringo being at the party that either Ringo or Elton John stays away; for it might be Elton John that stays away. Nevertheless, if someone runs in with the news that Ringo or Elton John might not be attending, we will hardly keep on assuming that all four Beatles will be there. Our shared information is weakened to: John, Paul, and George will be at the party.

¹⁷ The counter-response is that we would take steps to avoid this disaster by scaling back to an information state consistent with Ringo's non-attendance. I agree that this is what we would do; the question is whether scaling back should be understood as a repair strategy we use when disaster threatens, or as part of "might"'s basic semantic functioning. See Fuhrmann 1999. Thanks here to Thony Gillies.

It seems from these examples that "might φ ," uttered in information state S, has or can have the effect of cutting S down to a weaker information state S'; and it can have this effect both when φ is consistent with S and when φ is inconsistent with S. If information states are modeled as sets of worlds, then the effect of "might φ " is to add on additional worlds. The question, of course, is which additional worlds. Our reason for looking at deontic modals is that the analogous question about them was raised years ago by David Lewis, in a paper called "A Problem about Permission."

Lewis starts by describing a simple language game. The players are Master, Slave, and Kibitzer, though we'll be ignoring Kibitzer (he's used to it). Master issues commands and permissions to Slave, thereby shrinking and expanding what Lewis calls the sphere of permissibility, the set of worlds where Slave behaves as he ought. Behaving as he ought is Slave's only purpose in this game, and given how we defined the sphere of permissibility, that comes to behaving so that the actual world lies within the sphere. Slave can't work to keep the actual world within the sphere, though, unless he knows where its boundaries lie. Let's try to help him with this: how does the sphere evolve over time?

When the game begins, let's assume, all worlds are in the sphere of permissibility. Now Master begins issuing commands and permissions. Our

job is to figure out the function that takes a given sequence of commands (written $!φ$) and permissions (written $iφ$) to set of worlds permissible after all those commands and permissions have been given. That fortunately boils down to two simpler-seeming sub-tasks: first, figure out the effect of a command on the sphere of permissibility; second, figure out the effect of a permission on the sphere of permissibility.

You might think the second sub-task would be the easier one: after all, a sphere of permissibility would seem to be more directly responsive to permissions than commands. But it is actually the first sub-task that's easier. Suppose the going sphere of permissibility is S and Master says, "Mop that floor!" Then presumably the new sphere S' is the old one S , restricted to worlds where the floor gets mopped. The rule stated generally is

$$!φ: S \rightarrow S \cap \llbracket φ \rrbracket,$$

or, to formulate it as an identity:

$$(C) \quad !φ(S) = S \cap \llbracket φ \rrbracket.$$

The left-to-right inclusion here ($!_{\varphi}(S) \subseteq S \cap \|\varphi\|$) follows from two extremely plausible assumptions:

(c1) commands shrink (i.e. don't expand) the sphere, and

(c2) commands to φ make all $\sim\varphi$ -worlds impermissible

The right-to-left inclusion ($!_{\varphi}(S) \supseteq S \cap \|\varphi\|$) follows from (c1) and a third plausible assumption

(c3) commands to φ make only $\sim\varphi$ -worlds impermissible.

All of this is treated by Lewis as relatively undebatable, and nothing will be said against it here; it serves as background to the problem to come.

That problem concerns permission. If commands go with intersection, the obvious first thought about permissions is that they would go with unions:

$$j_{\varphi}: S \rightarrow S \cup \|\varphi\|,$$

or, the corresponding identity,

$$(P) \quad i\varphi(S) = S \cup \|\varphi\|.$$

The left to right inclusion ($i\varphi(S) \subseteq S \cup \|\varphi\|$) is hard to argue with; it follows from

(p1) permissions expand (i.e., do not shrink) the sphere, and

(p2) permission to φ renders only φ -worlds permissible.

But the right to left direction requires along with (p1) the principle

(??) permission to φ renders all φ -worlds permissible.

And while it is hard to argue with

(p3) permission to φ renders some φ -worlds permissible,

(??) seems clearly wrong. Lewis explains why:

Suppose the Slave had been commanded to carry rocks every day of the week, but on Thursday the Master relents and says to the Slave, 'The Slave does no work tomorrow'....He has thereby permitted a holiday, but not just any possible sort of holiday...[not] a holiday that starts on Friday and goes on through Saturday, or a holiday spent guzzling in his wine cellar...(27).

So (??) allows in too much. (p3) on the other hand, although correct, can't be the whole story. Not any old expanded sphere that contains φ -worlds will do--for the sphere whose sole φ -world has Slave staying on holiday through Saturday won't do. The situation is this:

Some worlds where the Slave does not work on Friday have been brought into permissibility, but not all of them. The Master has not said which ones. He did not need to; somehow, that is understood (27).

If it is understood, there must be a way we understand it: there must be a rule or principle of sphere-evolution that captures our shared implicit understanding of how permissions work.

Now we reach the problem of Lewis's paper. What is that rule? Or, to put it negatively, what exactly is wrong with a rule R that tells us that having been permitted to take Friday off, Slave can take that and other days off? Lewis looks at five answers.

(1) *R lets in **more worlds** than necessary*

Putting in a Saturday-off world enlarges the sphere more than necessary to allow Friday-off worlds. It's a "gratuitous enlargement" in the sense of adding more worlds than necessary.

Lewis replies that any reasonable enlargement will be gratuitous in that sense, since the only non-gratuitous enlargement adds in just a single world. This is fair enough, but it is not, I think, the "real" problem. If it were, then limiting ourselves to non-gratuitous (single-world) enlargements would address it. And it doesn't; for we could pick as our single world a world where Slave takes Saturday off too.

(2) *R lets in worlds **more remote** than necessary*

Putting the Saturday-off world in is a gratuitous enlargement in a qualitative sense. We should allow in only the closest worlds where the permitted action is done.

This, Lewis says, is too restrictive. Suppose Slave had been ordered to carry rocks around. Then he is forced to spend his vacation lifting weights! For weight-lifting worlds are closer to rock-carrying worlds than lying around-at-the-beach worlds are to rock-carrying worlds.

One can put the problem like this. A permission should cleanly cancel relevant earlier commands; but on the present approach supposedly cancelled commands continue, from beyond the grave as it were, to exert an effect. The clean cancellation requirement, as I will call it, will come up again.

(3) *R lets in worlds **more impermissible** than necessary*

Putting the Saturday-off world in is a gratuitous enlargement not in a qualitative but a prescriptive sense. We should put in the least impermissible worlds where the permitted action is done. Taking Friday and Saturday off was more impermissible than taking Friday off, so the two-day-off worlds aren't added in.

The objection Lewis offers is that this "solution" just restates, indeed aggravates, the problem: figuring out how comparative impermissibility evolves under the impact of commands and permissions is no easier (and possibly harder) than figuring out how straight permissibility does.

But there is again a prior worry – a version of the clean cancellation problem. Suppose Master first says not to eat any animals, then relents and permits eating lobster. Before lobster-eating was permitted, it was less impermissible to nibble on lobster than to eat a lobster in its entirety. So afterward is it only permissible to nibble on lobster?

(4) *R lets in worlds more disagreeable to Master than necessary*

Putting the Saturday-off world frustrates Master's known or guessable purposes.

Lewis objects that either Slave knows Master's purposes or he doesn't. If he does, there's no need for commands; he can work unsupervised. If he doesn't, then the principle cannot be what's guiding him.

Once again, there is a prior worry. Let's say that Master ordered Slave to carry rocks up the hill. Presumably she did this because she wants the rocks up the hill. But the Friday-off worlds that best serve the purpose of getting the rocks up the hill are ones where Slave invites his friends to play a game where two teams compete to see who can carry more rocks up the hill. This is again a version of the clean cancellation problem.

(5) *R lets in worlds violating more commands than necessary*

This takes a bit more explanation. It's a given that Master doesn't issue commands and permissions unless she needs to. She doesn't issue the command to φ if it is already impermissible for Slave not to φ ; and she doesn't issue permission to φ if Slave is already permitted to φ . In particular, then, Master would not have permitted Slave to take Friday off unless taking Friday off would otherwise have been an act of disobedience, an act in violation of some explicit or understood command. So, proposal: the effect of permitting φ should be to invalidate any commands that forbid φ ing – that are inconsistent with φ – while leaving other commands in place. The problem with an update rule that lets Saturday-off worlds into the sphere is that it invalidates more commands than necessary. To make Slave's taking Friday off permissible, it is enough to invalidate the work-Friday command; the work-Saturday command doesn't care if Slave takes Friday off, so it should be left in place.

Call this the remainder rule, because it defines S^+ as the set of worlds satisfying the commands that remain when all φ -inconsistent commands are invalidated. Lewis doesn't like the remainder rule either; here is why. Clearly, to apply the rule, we need there to be a list of commands ψ_1, \dots, ψ_k such that a world is permissible iff it complies with all of them, that is,

$$S = \|\psi_1\| \cap \|\psi_2\| \cap \dots \cap \|\psi_k\|.$$

For the way the rule works is we delete from this list all the ψ_i s inconsistent with φ , and let the commands that remain define S^+ . So if the φ -incompatible commands are $\psi_{j+1}, \psi_{j+2}, \dots, \psi_k$, the new sphere is

$$S^+ = \|\psi_1\| \cap \|\psi_2\| \cap \dots \cap \|\psi_j\|.$$

Where is the initial set of commands supposed to come from, though, the one we thin out to arrive at the reduced command-set that defines the S^+ -worlds? It would be one thing if " $i\varphi$ " were the first permission uttered; for then Master's earlier utterances were all commands, and we can let the ψ_i s be those commands. Ordinarily, though, " $i\varphi$ " is preceded by commands and other permissions. One could try considering just the commands that have already been given, ignoring the permissions, but these will not define

the current sphere of permissibility, because the update effects of earlier permissions will have been ignored.

It seems, then, that we are driven to contriving, reverse-engineering if you like, a package of commands that define the current sphere, a set of ψ_i s that together define S. Unfortunately the relation between S and packages of commands defining S is one-many; lots of them will issue in the same sphere of permissibility. How does Slave know which package to use? It makes a difference, because the effect on S of permitting φ varies enormously with our choice of implicit commands ψ_i .

So, for instance, suppose that the current sphere S = the worlds where Slave works all day, every day from Monday to Sunday; but we arrived at that sphere by a complicated series of enlargements and contractions that offers no clues to what the right ψ_i s, the right implicit commands, are. Slave might think that initially, before he is given Friday off, the commands in effect are

ψ_1 : Slave carries rocks on Monday.

ψ_2 : Slave carries rocks on Tuesday.

ψ_3 : Slave carries rocks on Wednesday.

ψ_4 : Slave carries rocks on Thursday.

ψ_5 : Slave carries rocks on Friday,

ψ_6 : Slave carries rocks on Saturday

ψ_7 : Slave carries rocks on Sunday.

The one command here inconsistent with “Slave takes Friday off” is “Slave carries rocks on Friday.” Suspending that one command leaves the commands to work other days still in place. Clearly on this way of doing it, Slave has not been permitted to take other days off, which was the desired result. But Slave might also think that the implicit commands are

χ_1 : Slave carries rocks on weekdays.

χ_2 : Slave carries rocks on the weekend.

Now the φ -inconsistent rule, the one to be cancelled on the present hypothesis, is “Slave carries rocks on weekdays.” But then the sphere of

permissibility expands to include all worlds where Slave works on the weekend. And that seems crazy. Master meant to give Slave Friday off, not Monday - Thursday as well!

Lewis's objection in a nutshell is that the implicit commands are too unconstrained for the remainder rule to be of any use. He may be right in the end. I wonder, though, whether there are constraints he is missing – constraints that don't come into view until you raise the problem he is pointing to in the starkest possible terms. Let's look, then, at the most extreme cases of badly chosen implicit commands. At the one extreme we have commands each of which is inconsistent with φ ; at the other we have commands none of which is inconsistent with φ . An example of each-inconsistent is

θ_1 : Slave carries rocks every morning of the week.

θ_2 : Slave carries rocks every afternoon of the week.

Neither of these is compatible with Slave taking Friday off. Cancelling the φ -inconsistent commands, then, is cancelling all commands whatsoever. If all commands are cancelled, then everything is permitted. Master wanted to let Slave take Friday off, but winds up giving Slave his freedom.

Now consider commands none of which individually requires Slave to work Friday, but whose joint effect is to require Slave to work every day of the week. For instance,

σ_1 : Slave carries rocks every morning if any afternoon.

σ_2 : Slave carries rocks every afternoon if every morning.

σ_3 : Slave carries rocks some afternoon.

σ_1 allows Slave to take Friday off, provided he never carries rocks in the afternoon. σ_2 allows him to take Friday off, provided he omits to carry rocks some morning. σ_3 allows him to take Friday off, provided he carries rocks some afternoon. Each of the σ_i s is consistent with φ , so none of them is cancelled on the present rule. But then the sphere of impermissibility never changes. Master tried to give Slave permission to take Friday off, but it turns out he still has to work on Friday.

What can we conclude from this? The remainder rule – the one that says to cancel all and only pre-existing commands that forbid φ – can give very silly results. But we can make that work in our favour, by letting the results' very silliness help us tighten the rule. Call a command-list reasonable if running it through the remainder rule yields an expansion satisfying (p1)-(p3) above.

(p1) permissions (expand) do not shrink the sphere, and

(p2) permission to φ renders only φ -worlds permissible.

(p3) permission to φ renders some φ -worlds permissible.

It is not hard to establish the following (proof in appendix):

FACT If S is defined by a reasonable list of commands, then $S = \|\sim\varphi\| \cap \|\psi\|$ for some ψ . Equivalently, any reasonable command list is of the form (up to equivalence) "You must not φ ," "You must ψ ."

This transforms the problem in a helpful way. Before we had one equation in several unknowns (corresponding to the several choices of implicit commands ψ_i). Now we have one equation in one unknown. For we know what S is; it's the present, pre-permission-to- φ , sphere of permissibility. And we know what $\|\sim\varphi\|$

is; the set of worlds where the permitted behaviour does not occur. The one unknown is $\|\psi\|$, that is, the new sphere of permissibility S^+ .

So, to review what we know. Whenever a permission to φ is issued, it's as though the initial command list had consisted of two commands:

first, one saying (precisely) do not φ .

second, a command ψ that allows φ ing,

Our job as sphere-redrawers is to throw out the do not φ command and form the set of worlds allowed by the command that remains. This is nothing like an algorithm, because there is more than one way of choosing the command ψ that remains. (There are many sets whose intersection with $\|\sim\varphi\|$ is S .) But it is instructive nevertheless.

One way the equation $S = \|\sim\varphi\| \cap \|\psi\|$ helps is by showing us how to conceive the task diagrammatically (see Figure 1). We are given

(i) the red region – the φ -worlds, the ones where Slave takes Friday off, as he has been permitted to do;

(ii) the blue region – the $\sim\varphi$ -worlds, the ones where Slave works Friday, in accord with his pre-permission obligations; and

(iii) the green region – the set of initially permissible worlds S , the ones where Slave works all week.

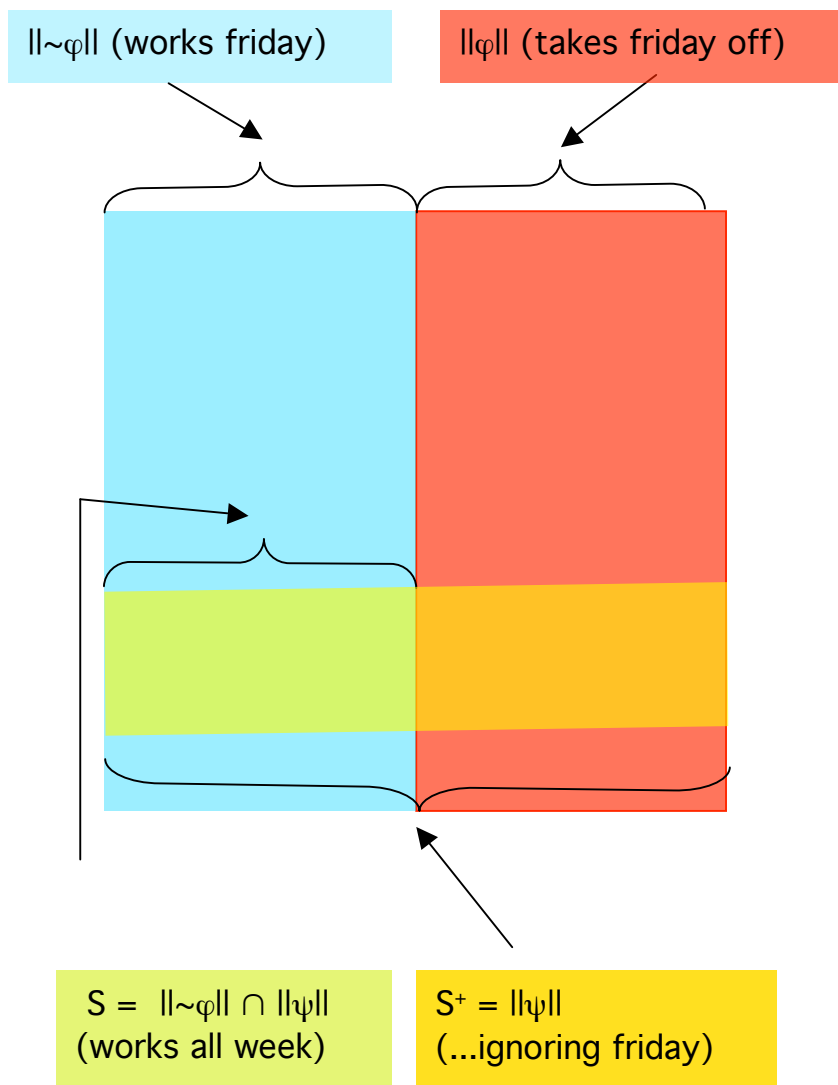


Figure 1

Our job is to extrapolate the green region beyond the bounds imposed by the blue region, thus arriving at the set $\|\psi\|$ of worlds that are permissible after Master cancels the command to work Friday.¹⁸

How is this to be done? The question here has nothing special to do with permission; it concerns logical extrapolation as such. To state it in full generality (see Figure 2): Suppose P is a proposition implying Q, represented by a subregion of the Q-region of logical space. When does a proposition R count as extrapolating P beyond Q to the rest of logical space?¹⁹

Here is a proposal about that, developed elsewhere. For R to extrapolate P beyond Q – for it to go on in the same way, as it were – R should meet three conditions:

- (i) within Q, R is true (false) in the same worlds w as P
- (ii) within Q, R is true (false) in w for the same reasons as w is a Q&P-world rather than a Q&~P-world (the reverse)
- (iii) outside Q, R is true (false) for the same reasons as within.

¹⁸ Another way to think of it is this: we are to decompose the green of the green region into its blue and yellow components, and allow the yellow component to unfold into the remainder of logical space – where, against a background of red now, it presents as orange rather than green.

¹⁹ “R” is meant to suggest remainder.

(i) says that R is equivalent to P within Q, so call it Equivalence. (ii) speaks to the reasons why R is true (false) within Q , so call it Reasons. (iii) takes a bit more explanation. For R to acquire new truthmakers (or

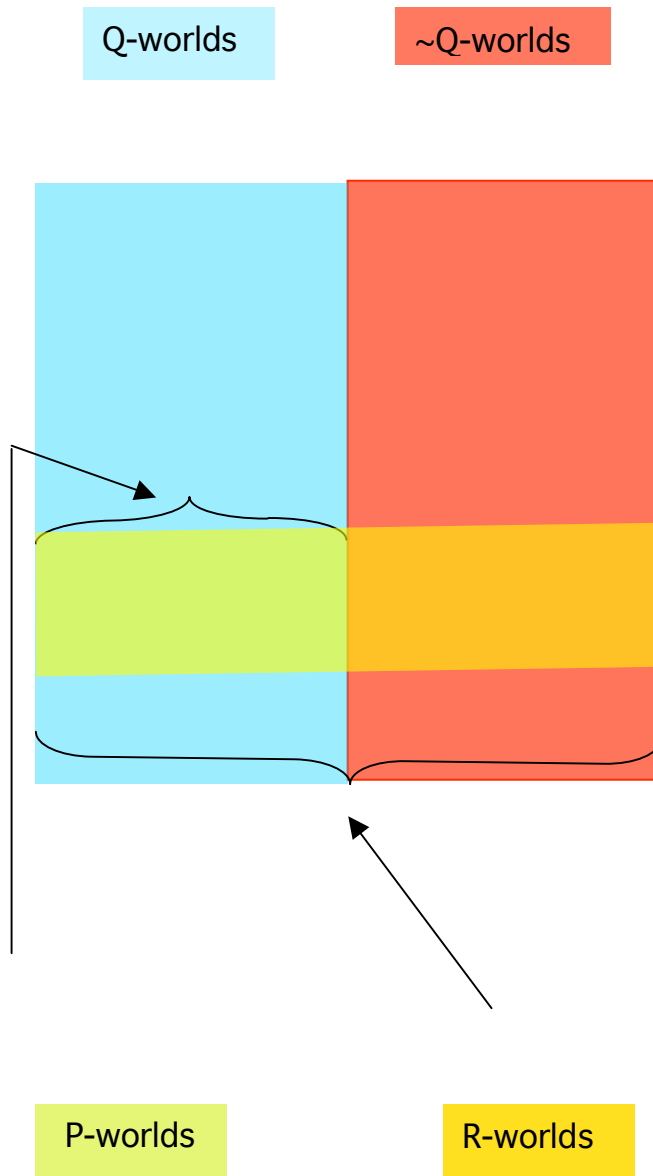


Figure 2

falsemakers) when it left the Q-region would mean that R was true (false) for one kind of reason in Q-worlds and another in non-Q-worlds. R would in that sense have changed direction as it crossed the Q-border; (iii) can thus be conceived as a kind of Orthogonality condition.

Now we construct a possible-worlds proposition P-Q that satisfies (i), (ii), and (iii).²⁰ As the notation suggests, P-Q can be thought of as the remainder when Q is subtracted from P. It suffices to say in which worlds P-Q is true and in which worlds it is false. Consider first the Q-worlds. Clearly P-Q's truth-value in any Q-world w must be the same as P's (by Equivalence). What is not so obvious is how to evaluate P-Q in worlds where Q is false (and therefore P, which implies Q, is false). The proposal: P-Q is false in a world where P and Q are both false if and only if P "adds falsity" to Q in the following sense:

X adds falsity to Y in w iff X&Y has a Y-compatible falsemaker in w.

X adds truth to Y in w iff its negation adds falsity there, that is, $\sim X \& Y$ has a Y-compatible falsemaker in w, which can equally be conceived as a Y-compatible truthmaker for $Y \supset X$:

X adds truth to Y in w iff $Y \supset X$ has a Y-compatible truthmaker in w

²⁰ (i) can be read off the definition. (ii) and (iii) depend on P-Q being assigned the appropriate truth- and falsemakers. Details are given elsewhere.

P-Q is true in w iff P adds truth to Q in w without adding falsity to Q in w.
P-Q is true, in other words, iff P adds truth and only truth to Q in w.
Assuming as above that P implies Q, this gives "-" the following quasi-truth-table:

P	Q	P-Q
t	t	t
t	f	impossible (P implies Q)
f	t	f
f	f	f

P adds falsity P adds only truth P adds nothing

The proposed update rule for permissions can now be stated very simply: *the effect of permitting φ is to subtract $\llsim\varphi\ll$ from the going sphere of permissibility.*

(UR_j) Assuming φ was initially impermissible, $S + j\varphi = S - \llsim\varphi\ll$.

How does the rule work in practice? How, for instance, do we find the worlds that are still impermissible after Master permits φ ? A world is still impermissible after Master permits φ iff $S - \llsim\varphi\ll$, which defines the new and enlarged bounds of permissibility, is false in that world. $S - \llsim\varphi\ll$ is false in w iff S is false in w for a reason not implying $\ll\varphi\ll$; equivalently, w is impermissible for a reason not implying $\ll\varphi\ll$. So a world is still impermissible iff it was impermissible for reasons additional to any violation of the ban on φ ing, reasons could still have obtained even if that ban had been observed.

Consider a world where Slave continues his holiday through the weekend. It was impermissible for reasons compatible with Slave's working Friday, such as his failure to show up on Saturday; so it remains impermissible after permission is given to take Friday off. A world where Slave takes just Friday off, however, does not remain impermissible, for it was impermissible only for φ -entailing reasons, and φ -entailing impermissibility-makers are deactivated when Master permits Slave to φ .

I said that a world where Slave takes just Friday off “does not remain impermissible.” Is that the same as becoming permissible? Not quite. An impermissible world is one where $S - \llsim\varphi\ll$ is false, and a permissible world is

one where $S \models \neg \varphi$ is true; and for a remainder-proposition to avoid falsity is not yet for it to achieve truth. $S \models \neg \varphi$ is non-false in a world w iff

- a) S adds no falsity to $\models \neg \varphi$ in w , that is, S has no $\models \neg \varphi$ -compatible falsemakers there,

For $S \models \neg \varphi$ to be true in w requires further that

- b) S adds truth to $\models \neg \varphi$ in w , that is, $\models \varphi \supset S$ has $\models \neg \varphi$ -compatible truthmakers there.

A world where Slave works Monday-Thursday and the weekend does indeed become permissible after Master permits Slave to take Friday off, because (a) and (b) are both satisfied, and for similar reasons. (a) is satisfied because Slave's working Monday-Thursday and the weekend means that $\models \text{Slave works all week}$ is false in w only because Slave took Friday off in it; and Slave's taking Friday off is obviously not compatible with his working Friday. (b) is satisfied because Slave's working Monday-Thursday and the weekend is a truthmaker for $\models \text{Slave works Friday} \supset \models \text{Slave works all week}$ that is compatible with Slave's working Friday. But although conditions (a) and (b) come to roughly the same in the case at hand, they are by no means equivalent.

Imagine that Master initially commands Slave to carry exactly one shovel and to keep it shiny clean. Permission is then given not to carry exactly one shovel. What should we say about a world u where Slave carries two shovels, one clean and one dirty? It meets condition (b) but not (a). It meets (b) because u does contain a truthmaker for $\models \text{Slave carries exactly one shovel} \supset \models \text{Slave carries exactly one shovel and keeps it clean}$ that is

compatible with Slave's carrying exactly one shovel, viz. the fact that Slave carries a clean shovel. It violates (a) because the fact that Slave also carries a dirty shovel is a \ll Slave carries exactly one shovel \ll -compatible falsemaker for \ll Slave carries exactly one shovel and keeps it clean \ll . A world v where Slave carries no shovels meets (a) but not (b). It meets (a) because the one and only reason \ll Slave carries exactly one shovel and keeps it clean \ll is false in v is that Slave carries no shovel there, and carrying no shovel is obviously not compatible with carrying exactly one shovel. It violates (b) because the one and only reason \ll Slave carries exactly one shovel $\ll \supset \ll$ Slave carries exactly one shovel and keeps it clean \ll is true in v is again that Slave carries no shovel there. The upshot is that while it is not impermissible for Slave to carry no shovels, it is not positively permissible either; the deontic status of a world where Slave carries no shovels is left indeterminate in this case. And while it is impermissible for Slave to carry two shovels, one dirty and one clean, it is not as thoroughly impermissible as carrying two dirty shovels, since S adds some truth in the first scenario, but it adds only falsity in the scenario where Slave keeps both of his shovels dirty.

Now I want to explore the epistemic analogue of Lewis's Master-Slave game. Here is how I understand the new game to work.

(1) The players this time are Teacher and Student, and the sphere of permissibility becomes the sphere of believability.

(2) The old game had Slave constantly adjusting his plans to fit with changes in what was permissible; the new one has Student constantly adjusting his theory to fit with changes in what is believable.

(3) It contracted the sphere of permissibility when Master said, "Do ψ "; the sphere expanded when Master said, "You may do φ ." Likewise it contracts the sphere of believability when Teacher says, " ψ is so"; the sphere expands when Teacher says, " φ might be so."

(4) There was no great mystery about the kind of contraction brought on by "Do ψ "; one simply rejected as impermissible worlds where ψ failed. Similarly there is no great mystery about the kind of contraction brought on by " ψ is so"; worlds where ψ fails are rejected as unbelievable.

(5) It was initially mysterious how "You may do φ " enlarged the sphere of permissibility. Similarly it is now mysterious to begin with how " φ may be so" enlarges the sphere of believability.

Let's continue the pretense that " φ might be so" has no effect on a sphere of believability that contains φ -worlds; it is only when all believable worlds are $\sim\varphi$ that we get an expansion. The question is, what expansion do we get? I propose that the update rule is pretty much as before.

(UR \diamond) Assuming φ was initially unbelievable, $S + \diamond\varphi = S - \|\sim\varphi\|$.

To see how it works, imagine that Teacher starts by saying it will be dry all week, meaning Monday-Sunday. She thereby banishes from the sphere of believability all worlds where it rains on one or more days. When Teacher learns that her evidence as regards Friday was shaky, she says, "Hold on, it might rain on Friday after all." Which worlds has she put back into the sphere of believability? To put it another way, what remains of Teacher's original prediction of no rain all week, once she has conceded it might rain on Friday? Our update rule says that

$$\begin{aligned} & \|\text{It will be dry all week}\| + \diamond(\text{It will rain on Friday}) \\ = & \|\text{It will be dry all week}\| - \|\text{It will be dry on Friday}\|. \end{aligned}$$

Now, $P \supset Q$ is non-false in w iff P adds no falsity to Q in w , and true iff P furthermore adds truth to Q in w ; and P adds falsity (truth) to Q in w iff P is false in w for a Q -compatible reason ($Q \supset P$ is true for a Q -compatible reason). Thus the worlds Teacher is representing as no longer unbelievable are those in which $\text{It will be dry all week}$ is false only because it rains on Friday, which is to say the worlds where it is dry on Monday, Tuesday, Wednesday, Thursday, Saturday, and Sunday. These worlds are also now believable, for its being dry on those days is a dry-Friday-compatible truthmaker for the conditional hypothesis that it will be dry all week if it is dry on Friday.

I want to return now to a limitation of the Lewis game noted earlier. Lewis stipulates that permission to φ has no impact unless φ was antecedently forbidden. That makes nonsense both of out-of-the-blue, discourse-initial, permissions, and permissions that soften commands that didn't strictly forbid the now-permitted behavior. Suppose, for instance, that Master, having first commanded Slave to work all week, allows him to go to the beach on his birthday; this wasn't strictly forbidden because Slave's birthday might still be a few months off. Similar worries can be raised about our epistemic analogue of the Lewis game. We have been assuming that "might φ " has no impact on the

sphere of believability unless φ was antecedently denied. But then what is going on in a conversation like this:

A: Where is Bob?

B: Hmmm, I don't really know, but he might be in his office.

[A: *I never said he wasn't.]

Or this:

A: Bob will be at the office tomorrow.

B: Not so fast, he said he might stay home on his birthday.

[A: *That's compatible!]

Call this the problem of unforced retractions. I see two ways of addressing it. The first is simply to strike “Assuming φ was initially impermissible,…” from the update rule for permissions, and to strike “Assuming φ was initially unbelievable…” from the update rule for “might.” These provisos might have been thought indispensable. For $S + \diamond\varphi = S - (\sim\varphi)$; and didn't our definition of P-Q assume that P implied Q?

In fact, it didn't. What is true that we were assuming that P implies Q when we gave the definition. But the definition itself makes sense either way. P-Q is the proposition that is false in w iff P&Q is false there for a Q-compatible reason, and true in w iff \sim P&Q is false there for a Q-compatible reason (and P&Q isn't). Nothing here depends on P implying Q.

I suspect this first response to the problem of unforced retractions is ultimately to be preferred. But let me here try a different response, one that is more in keeping with Lewis's approach in "A Problem About Permission." Lewis confronts there something like the dual of the present difficulty. Having laid it down that commands can only shrink the sphere of permissibility, he remarks that

One sort of commanding may seem to require special treatment: commanding the impermissible. Suppose that $\|\varphi\|$ contains no worlds that are ...permissible...The Master may nevertheless wish to command ...that φ ...Having commanded at dawn that the Slave devote his energies all day to carrying rocks, the Master may decide at noon that it would be better to have the Slave spend the afternoon on some lighter or more urgent task. If the master simply commands...that φ , then no worldremains permissible; the Slave, through no fault of his own, has no way to play his part by trying to see to it that the world remains permissible...Should we therefore say that in this case the sphere evolves not by intersection but in some more complicated way? I think not....What the Master should have done was first to permit and then to command that φ (27).

He notes a possible fix: whenever φ is impermissible, "a command that φ is deemed to be preceded by a tacit permission that φ , and the sphere of permissibility evolves accordingly" (27). Our present concern can be put in similar language:

One sort of permitting may seem to require special treatment: permitting the not impermissible. Suppose that the sphere of permissibility contains φ -worlds...The Master may nevertheless wish to permit...that φ ... Having at dawn permitted the Slave to take the day off, the Master may decide at noon that the Slave should be permitted to visit his mother this week. If the Master simply permits...the Slave to visit his mother this week, then no additional worlds...become permissible; for there are already permissible worlds where the Slave visits his mother, namely worlds where the Slave visits his mother today...Should we therefore say that in this case the sphere evolves not by the remainder rule but in some more complicated way?

I propose to avoid saying this by a maneuver similar to Lewis's: whenever φ is already permissible, permission to φ is deemed to be preceded by a tacit command not to φ , with the sphere of permissibility evolving accordingly. Likewise whenever φ is already believable, "it might be that φ " is imagined to be in response to the unspoken assertion that $\sim\varphi$.

How much justice does this kind of maneuver do to our feeling of still conveying something when we permit the not previously permissible, or suggest that things might be a way that no one had ever said they weren't?

The first thing to notice is that, just as permitting and then immediately commanding that φ can (even by our existing rules) change the sphere of permissibility, forbidding and then permitting φ can change the sphere of permissibility too. Mathematically speaking, there is no reason whatever to expect that $S^+ = \text{j}\varphi(\text{!}\sim\varphi(S)) = (S \cap \|\sim\varphi\|) - \|\sim\varphi\|$ will just be S again. Indeed there is reason to expect it often won't; for the equation

$$(*) \quad (P \ \& \ Q) - Q = P$$

fails when P and Q overlap.²¹ Suppose, to start with a case of total overlap, that P and Q are one and the same proposition. Then $(P \ \& \ Q) - Q = (P \ \& \ P) - P = P - P$. $P - P$ is presumably the null proposition; anyway it is not P as $(*)$ would require. Or let P be the conjunctive proposition $Q \ \& \ R$, where R is thoroughly independent of Q . $((Q \ \& \ R) \ \& \ Q) - Q$ should be $Q \ \& \ R$, according to $(*)$. But we know by Boolean algebra that $(Q \ \& \ R) \ \& \ Q$ is $Q \ \& \ R$; and the result of subtracting Q from $Q \ \& \ R$ is not going to be $Q \ \& \ R$ again.

In principle, then, forbidding and then permitting φ can change the sphere of permissibility. Here is an example where it happens. Imagine that Master

²¹ I write “ $P \ \& \ Q$ ” instead of “ $P \ \cap \ Q$ ” to remind us that $P \ \cap \ Q$ considered as a proposition is the conjunction of P and Q .

starts out by commanding Slave to work on her (Master's) birthday. The initial sphere S is thus the set of worlds where Slave works on Master's birthday. Then Master further commands that Slave is to work on Friday. $S + !\sim\varphi = S \cap \|\sim\varphi\| =$ the set of worlds where Slave works on Master's birthday and on Friday. Then Master permits Slave not to work on Friday after all. The resulting sphere $(S \cap \|\sim\varphi\|) - \|\sim\varphi\|$ is not $S =$ the set of worlds where Slave works on Master's birthday, but $S^* =$ the set of worlds where Slave works on Master's birthday unless it falls on a Friday. (The added worlds were impermissible only because Slave took Friday off in them, when Friday was Master's birthday; worlds that were impermissible only because they violate a certain ban do not remain so when the ban is lifted.) Forbidding and then immediately permitting Slave to take Friday off can thus have a non-trivial effect on the sphere of permissibility. Like remarks apply to asserting it will rain on Friday and then immediately taking it back, having previously asserted that it will rain on Teacher's birthday,

I have argued that forbidding and then immediately permitting φ can change the sphere of permissibility, and also that asserting φ and then immediately allowing that maybe $\sim\varphi$ can change the sphere of believability. But there are also cases where permitting what I've just forbidden (admitting that a previous assertion might be wrong) leaves the sphere just as it was. An example might be this. Nothing has been said about Bob's location, but I know you want to find him.

What is accomplished by saying "He might be in his office," when no one has suggested otherwise? Likewise what is accomplished by announcing out of the blue that it is permitted to climb trees in order to rescue kittens?

It seems to me these things are not so mysterious, once we distinguish what has been forbidden, in the sense that the command has been given, and what is forbidden, in the sense that it's against the rules but Counselor may not have got around to announcing it yet. The children may know to begin with that nothing has been forbidden, but they have no idea what might or might not be forbidden in the sense of being off limits or against the rules. When they hear that tree climbing after kittens is permitted, they learn an upper bound on what is forbidden, namely that it doesn't include tree climbing after kittens. This is not, to my mind, because the counselor has said climbing after kittens is not forbidden; she has the ability to forbid but not, as we're imagining the game, the ability to comment on the extent of the forbidden. What Counselor has done is "shown" that climbing after kittens is not forbidden by staging a confrontation with an imagined off-screen forbiddler, and canceling that imagined person's decree.

Something similar is going on when I tell the seeker after Bob that he might be in his office. The distinction we need this time is between what has been asserted, and what is understood to be so even if no one has got around to announcing it yet. Before I spoke, my friend might have been wondering whether an assertion that Bob was not in his office was in the cards. I satisfy her curiosity not by saying that an assertion to that effect is not in the cards; my subject matter is

Bob and his office, not assertions about them. I satisfy my friend's curiosity by showing that an assertion to that effect is not in the cards, by staging a confrontation with someone imagined to have made the assertion, and undoing what they are imagined to have done.

Let's return now to some of the problems raised at the outset, starting with problems for the standard semantics (SS). One problem was that SS gave “might φ ” the wrong subject matter. “Bob might be in his office” seems intuitively to be about whatever “Bob is in his office” is about. Neither concerns the speaker or the extent of her knowledge. The present view construes “might φ ” as a device for retracting or canceling an assertion of φ . φ has the same subject matter qua retracted as it has qua asserted.

A second worry was that the truth-conditions assigned by SS, in its naïve Moorean version, were too weak. If “might φ ” says only that my information doesn't rule φ out, why do I accept correction by observers with information that I did not possess?²² It is indeed puzzling why the speaker should accept correction, if the “mistake” was to have misstated the epistemic facts. But suppose that “might φ ” is not a statement of fact. Suppose it is a cancellation

²² See however the “Mastermind” example in von Fintel and Gillies 2007.

order, an attempt to undo or block the assertion that $\sim\varphi$, to expel or bar $\sim\varphi$ from the common ground. If that is what the better-informed observer is taking issue with, then her objection makes perfect sense; for however well-intentioned, the cancellation order was unfortunate. The observer knew that $\sim\varphi$, hence that it would have been better not to block $\sim\varphi$'s addition to the common ground.

As already discussed, the standard semanticist's response to the "too weak" objection is to make the truth-conditions stronger: φ should be consistent not only with my information, but all pertinent information – where the test of pertinence must presumably be that the speaker concedes that he was in error if that information really does obtain. The third worry was that these revised truth-conditions are too strong. If "might φ " is false when φ is ruled out by pertinent facts, then speakers should restrain themselves except when such facts are known not to obtain. I shouldn't say that Bob might be in his office, if there is a chance that Bob has unbeknownst to me been seen elsewhere. Clearly, though, speakers do not restrain themselves in this way. (Or "might"-claims would hardly ever get made.) Why not? Well, how careful I need to be depends on the threat I'm confronting. If I am concerned that my claim might be false, then I should indeed hold back until I have tracked down all the pertinent facts. But what if falsity is not the issue? What if I am concerned rather that my claim will turn out to have been ill advised, or counterproductive, given the purposes of the conversation? A blocking order is ill advised just to the extent that the assertions it blocks are better-informed than one's attempt to block them. The abstract possibility of evidence against φ somewhere out there does nothing to suggest my utterance of "might φ " is ill advised in

this sense; it does nothing to suggest that I am frustrating the efforts of better-informed co-inquirers. And the fact is that I do restrain myself when I run the risk of blocking better-informed assertions to the contrary. You will not hear me telling geophysicists that Mt St Helens might be about to blow.

The fourth problem was that SS is too epistemic. It reckons "I might vote for Kucinich and I might not" false unless I am genuinely undecided how I am going to vote. The present theory can say that I am showing my audience, by example as it were, that no assertion is to be expected on the topic of how I am going to vote. I do it by giving myself the opportunity to make that sort of assertion and then visibly passing it up.

The fifth problem was that "It might rain on Tuesday" does not seem entail "It might rain on Tuesday or Wednesday," even though a disjunction is compatible with the relevant body of information if one of its disjuncts. If anything the implication goes the other way: "It might rain on Tuesday or Wednesday" makes a stronger claim than "It might rain on Tuesday."²³ How are we to make sense of this? A "stronger claim" in the context of the cancellation theory is a might-

²³ The analogous phenomenon with permission is perhaps better known. Suppose you are hungry and I tell you: You may have a piece of cake or a piece of pie. You reach for the pie and I snatch it away. What gave you the idea that that was a permissible disjunct?

claim that cancels more. To show that $\diamond(\varphi \vee \psi)$ is stronger than $\diamond\varphi$, we show that whereas $\diamond\varphi$ cancels only $\sim\varphi$, $\diamond(\varphi \vee \psi)$ cancels $\sim\varphi$ and $\sim\psi$ both.

Teacher starts out by telling us that it will not rain at all this week. This initializes the sphere of believability to $S = \sim M \& \sim T \& \sim W \& \sim R \& \sim F \& \sim A \& \sim U$.²⁴ The weather report then leads her to qualify this claim: it might rain on Tuesday or Wednesday. By (UR_i) , the sphere now expands to

$$\begin{aligned} & \sim M \& \sim T \& \sim W \& \sim R \& \sim F \& \sim A \& \sim U + \diamond(T \vee W) \\ = & \sim M \& \sim T \& \sim W \& \sim R \& \sim F \& \sim A \& \sim U - \sim(T \vee W) \\ = & \sim M \& \sim T \& \sim W \& \sim R \& \sim F \& \sim A \& \sim U - \sim T \& \sim W. \end{aligned}$$

Now, a conjunction minus the conjunction of some of its conjuncts is surely the conjunction its other conjuncts.²⁵ So the remainder here is $\sim M \& \sim R \& \sim F \& \sim A \& \sim U$ —which is the result we were hoping for. But let us follow the calculation through:

²⁴ M, T,..., A, and U are the propositions that it rains on Monday, that it rains on Tuesday,..., that it rains on Saturday, and that it rains on Sunday.

²⁵ At least when the conjuncts are suitably disjoint, as the weather on one day is disjoint from the weather on another. I discuss content-parts and content-overlap elsewhere.

$\sim M \& \sim T \& \sim W \& \dots \& \sim U - \sim T \& \sim W.$
 $= \sim M \& \sim T \& \sim W \dots \& \sim U$
 $\cup \{w: \sim M \& \sim T \& \sim W \& \dots \& \sim U \text{ adds no falsity to } \sim T \& \sim W \text{ in } w\}^{26}$
 $= \sim M \& \sim T \& \sim W \& \dots \& \sim U$
 $\cup \{w: \sim M \& \sim T \& \sim W \& \dots \& \sim U \text{ lacks } (\sim T \& \sim W)\text{-compatible falsemakers in } w\}$
 $= \sim M \& \sim T \& \sim W \& \dots \& \sim U$
 $\cup \{w: \text{all } \sim M \& \sim T \& \sim W \& \dots \& \sim U \text{'s falsemakers in } w \text{ imply } T \vee W\}$
 $= \text{the worlds where it's dry all week}$
 $\text{plus the worlds where it isn't only because it rains Tuesday or Wednesday}$
 $= \text{the worlds where it's dry Monday and Thursday-Sunday.}$

Consider, for instance, a world x where it rains on Tuesday alone. x remains unbelievable iff among the reasons it was unbelievable are some that do not imply $T \vee W$. But the reasons x was unbelievable are one and all T -implying; they were entirely to do with its raining on Tuesday in x . Reasons that imply T are trivially reasons that imply $T \vee W$. So x is not unbelievable any longer. The same applies to worlds y where it rains just on Wednesday, and worlds z where it rains Tuesday and Wednesday; these worlds too were unbelievable only for $(T \vee W)$ -implying reasons, and that kind of reason is irrelevant now that we've learned

²⁶ I assume for simplicity that if $S + \diamond(T \vee W)$ is not false in a world, it is true there. Strictly speaking, $S \cap \{w : S \text{ adds no falsity to } \sim(T \vee W) \text{ in } w\}$ is the set of worlds where $S + \diamond(T \vee W)$ is not false; for truth, a little more is required

TvW might be true. Thus the effect of $\diamond(TvW)$ is to cancel the ban on worlds where it rains on Tuesday and/or Wednesday. The effect of $\diamond T$ followed by $\diamond W$ is the same. This accounts for the feeling that "It might rain on Tuesday or Wednesday" implies both that it might rain on Tuesday and that it might rain on Wednesday.

The sixth problem we raised for SS was the Yalcin problem: it has trouble explaining the incoherence of " φ & it might be that $\sim\varphi$." The problem isn't unassertability, for unassertable hypotheses can still be hypothesized, say, in the antecedent of a conditional. And it makes no sense to say, "If it rained last night, but it might not have rained, then the clothes we hung up will be wet." This is a problem, Yalcin contends, for any truth-conditional theory of "might." For the following is an invalid argument: It might be the case that $\sim\varphi$, therefore $\sim\varphi$. A one-premise argument "X, therefore Y" is invalid, one would think, only if there is an intelligible scenario where Y is false even though X is true. But a scenario where it is false that $\sim\varphi$ but true that $\sim\varphi$ might be so is a scenario where φ but maybe not---the very thing we have called unintelligible.

Recall that $\diamond\sim\varphi$ on the cancellation view is not²⁷ a device for stating facts; it makes little sense, then, to ask whether it can happen that $\diamond\sim\varphi$ is true to the facts without $\sim\varphi$'s being true to the facts. Instead of asking whether $\diamond\sim\varphi$'s truth forces $\sim\varphi$ to be true, we should ask whether cancelling the

²⁷ Not in the first instance, anyway.

assertion that φ commits one to asserting that $\sim\varphi$. It obviously doesn't. No wonder the argument " $\diamond\sim\varphi$, therefore $\sim\varphi$ " strikes us as invalid; accepting the premise puts one under no rational pressure to accept the conclusion. One question remains: why does " $\varphi \& \diamond\sim\varphi$ " seem incoherent, even as a supposition? The problem is not that no world can answer both to the specification that φ and the specification that $\diamond\sim\varphi$. It's that no world-specification can both demand that w be φ and fail to demand that w be φ . " $\varphi \& \diamond\sim\varphi$ " is suppositionally incoherent because it gives the would-be supposer contradictory instructions: they are to suppose that φ while at the same time taking care not to suppose that φ .

Update semantics, we said, left it un- or under-explained why allowing that Ringo might not go to the party leaves intact the information that John, Paul, and George will be there. The present theory says that a world w remains unbelievable only if "John, Paul, and George will be there" adds falsity in that world to "Ringo will be there." "John, Paul, and George will be there" adds falsity to "Ringo will be there" in w iff it is false in w for a reason compatible with Ringo's presence at the party. "John, Paul, and George will be there" is false in w for a reason compatible with Ringo's presence at the party iff John, Paul, or George misses the party in w . But then worlds where any of John, Paul, or George misses the party are still unbelievable after we learn that Ringo might not attend—which was the desired result. This is an instance of the problem of clean cancellation. Assertive content that entails the falsity of what we learned might be true is cancelled; the rest of what was asserted remains in place.

Our second worry about update semantics was this. It tells us that \diamond_{φ} uttered in information state S has no effect unless S and φ are inconsistent. Suppose, for instance, that $S = \{\sim\chi\}$ and $\varphi = \chi \vee \psi$. $\chi \vee \psi$ is consistent with $\sim\chi$, so update semantics says that $S + \diamond_{\varphi}$ should be S again. But that seems wrong. For recall that $\diamond(\chi \text{ or } \psi)$ has the same force as $\diamond\chi$ followed by $\diamond\psi$. To be told inter alia that $\diamond\chi$ presumably cancels the information initially present in $S = \|\sim\chi\|$. Here, then, is a case where $S + \diamond_{\varphi}$ is a proper subset of S , even though φ is fully consistent with S . What does the cancellation theory say? When φ is consistent with S , we imagine it preceded by an assertion of $\sim\varphi$. $S + \diamond_{\varphi}$ is $(S \cap \|\sim\varphi\|) - \|\sim\varphi\|$. If the two-part operation of asserting and then withdrawing $\sim\varphi$ always left S unchanged, then this maneuver would not gain us much. But we have seen that $(S \cap \|\sim\varphi\|) - \|\sim\varphi\|$ is not always S . And in the present case $(S \cap \|\sim\varphi\|) - \|\sim\varphi\|$ would seem to be a much weaker proposition than S .²⁸

That completes my explanation and defense of the cancellation theory. Many important topics have been left undiscussed; some of them I have thought about, others not. For instance,

- (a) Master-Slave and Teacher-Student are highly unnatural games. How does the sphere evolve when deontic/epistemic authority does not rest with one person?²⁹

²⁸ $\|\sim\chi\| \cap \|\sim(\chi \vee \psi)\| - \|\sim(\chi \vee \psi)\| = \|\sim(\chi \vee \psi)\| - \|\sim(\chi \vee \psi)\| = \|\sim\varphi\| - \|\sim\varphi\|$.

²⁹ Thanks here to Sally Haslanger.

(b) The focus has been on stand-alone "might"-statements. How should we understand "might"'s contribution in conditionals, or in complex predicates?

(c) If "might $\sim\varphi$ " is a device for rejecting an imagined assertion of φ , what is "must φ "? A device for rejecting an imagined instance of "might $\sim\varphi$ " –that is, for rejecting φ 's rejection? Can I reject φ 's rejection without going so far as to assert that φ ?³⁰

(d) How does the context-change function given in (UR \diamond) relate to the contraction operation studied in the literature on belief revision?³¹

I would like to take the opportunity in closing to cancel any would-be assertions to the effect that these topics will not be discussed in future work.

Proof of FACT

³⁰ This could bear on the puzzling weakness of "must φ " vis a vis φ .

³¹ Fuhrman 1996

From (p3), which says that the right expansion should bring in at least one φ -world, we conclude that any package of commands all of whose members are consistent with φ is unreasonable; for such a package fails to enlarge the sphere of permissibility, as it has to be enlarged to make room for φ -worlds. From (p2), which says that the right expansion should bring in only φ -worlds, we conclude that any package of commands none of whose members is consistent with φ is unreasonable. For that kind of package expands the sphere of possibility to include every world, and we know by (p2) that permission to φ should bring in only φ -worlds. So, any reasonable package of commands $\langle \psi_i \rangle$ has members consistent with φ and members inconsistent with φ . Let's use χ for the conjunction of all ψ_i s inconsistent with φ , and ψ for the conjunction all ψ_i s individually consistent with φ . Then

$S =$ the set of $(\chi \wedge \psi)$ -worlds

$S^+ =$ the set of ψ -worlds

$S^+ - S =$ the set of $(\psi \wedge \sim\chi)$ -worlds

Again, $\langle \psi_i \rangle$ is reasonable only if

(\exists) $S^+ - S$ contains φ -worlds (from (p3))

(\forall) $S^+ - S$ contains only φ -worlds (from (p2))

From (\exists) we learn that ψ is consistent with φ . Proof: Suppose not. Then S^+ (= the set of ψ -worlds) does not contain any φ -worlds. But (\exists) implies that S^+ does contain φ -worlds, since $S^+ - S$ contains them. A more interesting result follows from (\forall) : for all S^+ -worlds w , χ holds in w iff φ does not hold in w . The "only if" direction is easy, since each of χ 's conjuncts is by definition inconsistent with φ . For the "if" direction, suppose that χ does not hold in w . w cannot be an S -world because S -worlds have to satisfy all the ψ_i s. But then w is in $S^+ - S$. And according to (\forall) , every world in $S^+ - S$ satisfies φ . So, the implicit commands suitable to serve as backdrop to a permission to φ must be divisible into two parts: $\chi = \sim\varphi$ = the part that forbids φ ing, and ψ = the part that allows φ ing. QED

References

- DeRose, K. 1991. Epistemic possibilities. *The Philosophical Review* 100 (4): 581-605.
- Egan, A. 2007. Epistemic modals, relativism and assertion. *Philosophical Studies*, 133(1):1-22.

- Egan, A., J. Hawthorne, and B. Weatherson. 2005. Epistemic modals in context. In Preyer, Gerhard & Peter, Georg (eds). *Contextualism in philosophy: Knowledge, Meaning, and Truth*.
- von Fintel, K., and A. S. Gillies. CIA Leaks. *The Philosophical Review* 117(1): 77-98
- von Fintel, K. and A. S. Gillies. 2007. An Opinionated Guide to Epistemic Modality. *Oxford Studies in Epistemology* 2.
- Fox, D. ms 2006. Free Choice and the Theory of Scalar Implicatures.
- Fuhrmann, A. 1996. *An essay on contraction*. University of Chicago Press.
- Fuhrmann, A. 1999. When hyperpropositions meet. *Journal of Philosophical Logic* 28: 559-574
- Gillies, A. S. 2004. New foundations for epistemic change. *Synthese* 138 (1): 1-48.
- Hacking, I. 1967. Possibility. *Philosophical Review* 76 (April): 143-168.
- Hudson, J. L. 1975. Logical subtraction. *Analysis* 35: 130-135.
- Humberstone, L. ms 1981. Logical subtraction: Problems and prospects.
- Humberstone, L. 2000. Parts and partitions. *Theoria* 66 (1): 41-82.
- Jaeger, R. A. 1973. Action and subtraction. *Philosophical Review* 82: 320-9
- Kamp, H. 1973. Free choice permission. *Proceedings of the Aristotelian Society* 74: 57-74.

- Kratzer, A. 1981. The notional category of modality. In *Words, worlds, and contexts: New approaches in word semantics.*, eds. H. Eikmeyer, H. Rieser, 38-74. Berlin: de Gruyter.
- Lewis, D. 1979. A problem about permission. As reprinted in Lewis 2000: 20-33
- Lewis, D. 2000. *Papers in ethics and social philosophy.* Cambridge studies in philosophy. Cambridge: Cambridge University Press.
- MacFarlane, J. ms 2006. Epistemic modals are assessment-sensitive.
<http://sophos.berkeley.edu/macfarlane/epistmod.pdf>
- MacFarlane, J. 2005. Making sense of relative truth. *Proceedings of the Aristotelian Society* 105 (1): 305.
- Moore, G.E. 1962. *Commonplace Book 1919-53.* George, Allen, and Unwin, London
- Stanley, J. Fallibilism and concessive knowledge attributions. *Analysis* 65 (2): 126-31
- Teller, P. 1972. Epistemic possibility. *Philosophia: Philosophical Quarterly of Israel* 2 (October): 303-320.
- Thomson, J.J, and A. Byrne. 2006. *Content and modality : Themes from the Philosophy of Robert Stalnaker.* Oxford: Clarendon Press
- Veltman, F. 1996. Defaults in update semantics. *Journal of Philosophical Logic* 25 (3): 221-261.

Yablo, S. 2006. Non-catastrophic presupposition failure. Thomson and Byrne (eds.) 2006: 164-190.

Yalcin, S. 2007. Epistemic Modals. *Mind*, 116(464): 983

Zimmermann, T. E. 2000. Free choice disjunction and epistemic possibility. *Natural Language Semantics* 8 (4): 255-290.