Putting ‘Ought’s Together

The logic of ‘ought’ is unruly, or seems so. Apparent counterexamples to many logically satisfying principles surfaced almost the moment deontic logic got going. My focus is one of those principles, an inference pattern called **Agglomeration**:

(Agglomeration:) ought ϕ, ought $\psi \vDash$ ought $(\phi \wedge \psi)$

Agglomeration is a multi-premise closure principle for ‘ought’s: when two things ought to be the case, so too ought their conjunction. Some in the deontic logic tradition, perhaps most notably Jackson (1985), think Agglomeration invalid. But the classic view, which treats ‘ought’ as a necessity operator and so validates it, has an even longer tradition and has been vigorously defended as recently as von Fintel (2012).

There is something right and something wrong in both traditions, I argue. The classic view is right that deontic ‘ought’s agglomerate. Nonetheless, Agglomeration is not valid, just as the heterodoxy says. This is because *epistemic* readings of ‘ought’ do not agglomerate.

To make sense of all the data, I give a new semantics for ‘ought’, where it quantifies over *propositions* in a contextually supplied ordering, not worlds: $\ulcorner$ought $\phi\urcorner$ says that *some best proposition* entails ϕ. To predict why Agglomeration fails in epistemic rather than deontic cases, I enrich this basic semantics in a few ways. I propose that ‘ought’ is *question-sensitive* in a novel way: the propositions it quantifies over are partial answers to some background question. These questions determine what count as the relevant *ways* for a proposition to be true. ‘Ought’ also requires that the best propositions in context are (pairwise) consistent: no proposition entails the negation of another. In this framework, deontic and epistemic

Agglomeration have different statuses because their underlying orderings have crucially different properties: propositions can be epistemically but not deontically worse than all the relevant ways for them to come about.

1 Epistemic 'Ought'

Deontic readings of 'ought' and 'should' spring most readily to mind. A sentence like

(1) John ought to arrive by ten. (That's when the meeting starts.)

is naturally heard as talking about John's obligations. But 'ought' and 'should' also have non-deontic readings. Suppose Jane thinks the bus left 30 minutes ago and it usually takes 40 minutes to her nearest bus stop. She might truly say

(2) The bus ought to/should arrive in 10 minutes.

Here Jane does not say it would be *good* for the bus to arrive in 10 minutes. She is making a *prediction* about when the bus will arrive. Call this an *epistemic* reading of 'ought'.[1]

Other things being equal, we should seek a unified account of the epistemic and deontic 'ought'. When we find a single modal word expressing a variety of modal flavours, especially when this occurs across languages, most likely the variation in flavour results from different kinds of parameters being supplied by context to a unique semantic entry.[2] The combination of flavours across languages is very unlikely to occur by chance.

2 Agglomeration

Now for an observation that pushes us in a different direction: epistemic and deontic 'ought's seem to have a different *logic*. Agglomeration looks good for deontic

[1]There is disagreement about whether such 'ought's are genuinely epistemic or, as Yalcin (2016) argues, are *normality* readings of 'ought'. I aim to be neutral here, but use probability orderings later for concreteness.

[2]See von Fintel and Iatridou (2008) for a survey of the cross-linguistic data here.

‘ought’s but not for epistemic ‘ought’s.

2.1 Epistemic agglomeration

Epistemic Agglomeration can fail for a set of propositions each of which we expect to be true, but whose conjunction we do not. (Note this has the same structure as Makinson (1965)’s *preface paradox*.)

Take the following example:

> **The Office** There are 26 workers, Alice, Bob, Carol, ..., and Zadie, in our office. On average, they all come into work 17 days out of 20.

All of these sentences sound true here:

(3) Alice should be in work today.

(4) Bob should be in work today.

(5) Carol should be in work today.

...

(6) Zadie should be in today.

But, when ‘should’ takes wide scope, it does not seem to follow that:

(7) Everyone should be in today.

This is contrary to what Agglomeration predicts. If (3) – (6) are true, then (7) should be too.

2.2 Deontic agglomeration

To argue that deontic Agglomeration is valid, I’ll note how natural it seems and then respond to some putative counterexamples.

Start with a simple example. Suppose that I ought to help Alice and that I ought to help Billy. It is hard *not* to conclude I ought to help Alice and help Billy. It is just eminently plausible that if there are two things, each of which I ought to do, then I ought to do *both*. And we can keep agglomerating, as we add further obligations.

If I should help Alice, Billy *and* Carol, or Alice, Billy, Carol *and* Daniel, then in each case I ought to help all of them. And so on. Short of good counterexamples, we should regard deontic Agglomeration as valid.

But some think there are good counterexamples to it.[3] Jackson (1985) provides the following:

> **Chariots.** Attila and Genghis are driving their chariots towards each other. If neither swerves, there will be a collision; if both swerve, there will be a worse collision . . . but if one swerves and the other does not, there will be no collision. Moreover if one swerves, the other will not because neither wants a collision. Unfortunately, it is also true to an even greater extent that neither wants to be 'chicken'; as a result what actually happens is that neither swerves and there is a collision.

Jackson says the following are true here:

(8) Attila ought to swerve.

(9) Genghis ought to swerve.

But the following does *not* sound true:

(10) Attila and Genghis both ought to swerve.

This would be contrary to what Agglomeration predicts.

The objection presupposes that 'ought' is univocal throughout. But is it? Famously, there is a distinction between 'ought's that say what particular agents should do, the *ought to do*; and 'ought's that say how the world would be if things went best, the *ought to be*. Once we distinguish between them, we see that the counterexample fails.

The true readings of the premises seem to involve the 'ought to do' sense. (8) is naturally heard as saying

(11) Attila should make it the case that Attila swerves.

[3] Dilemmas, which are also regarded as a source of counterexamples to deontic Agglomeration, are addressed separately in section 7.

and (9) is naturally heard as saying

(12) Genghis should make it the case that he swerves.

So (8) and (9) do not involve the same *ought to do*. To bring this out notice that both of the following sound *false*:

(13) Attila ought to make it the case that Genghis swerves.

(14) Genghis ought to make it the case that Attila swerves.

This suggests an equivocation in the counterexample, when read as one about the *ought to do*. For there can be many *ought to do*s even in a given case. In ours, there is what *Attila* ought to do and what *Genghis* ought to do. But on neither notion are *both* of the premises true.[4]

When we move to an 'ought to be', then the premises are no longer clearly true. Neither of

(15) It ought to be that Attila swerves.

(16) It ought to be that Genghis swerves

seem true here. While (exactly) one of them should swerve, it does not matter which one: things go best if (just) Attila swerves or (just) Genghis swerves. So, on the *ought to be*, the counterexample falters too.

But then what 'ought's are true, when ϕ and ψ-ing are clearly jointly best, but $\phi \wedge \psi$-ing is not? *Pace* Jackson, I say Agglomeration does not fail in those cases. Instead, a principle I call *Indifference* holds:

> **Indifference.** In cases where ϕ_1, ... ,ϕ_n are jointly (intuitively) deontically best, $\ulcorner$ought $\phi_1 \vee ... \vee \phi_n\urcorner$ is true.

To bolster this, consider two kinds of cases. First, take cases where ϕ-ing and ψ-ing are equally good and better than anything else, but $\phi \wedge \psi$-ing is worse than doing

[4] What if in the *ought to do* 'ought' is syntactically a control verb, combining not with a sentence but a predicate and a name? This does not save the counterexample: on such a syntax, the natural way to formulate Agglomeration keeps the name fixed. (Especially if, as seems plausible, type-shifting should yield the semantics for the control verb.)

either individually, like:

> **Dessert**. I have three options for dessert: cannoli, cheesecake, and apple pie. The pie and the cannoli are the tastiest. I can order as many dishes as I like, but I will definitely feel ill if I have more than one dessert.

Here I should think:

(17) I ought to have pie or cannoli.

Second, take cases where ϕ-ing and ψ-ing are equally good, but contextually inconsistent, like:

> **Going Home**. There are three different routes home, A, B and C. A and B get us home equally quickly, but C takes longer. All we care about is how quickly we get home.

Here I should say

(18) I ought to take route A or route B.

In general, multiple best options make for disjunctive 'ought's.

2.3 Context-shifting in the Office?

We appealed to context-shifting in cases like **Chariots**. But couldn't there be context-shifting in **the Office** too?

The appeal to context-shifting fares worse here. For one thing, it is not obvious what to attribute it to. In **Chariots** we relied on the *agent*-sensitivity, of the *ought to do*. But the epistemic 'ought' is not sensitive to the subject of its prejacent.

There is also a more general reason to think that context-shifting will not help. Suppose context does shift and that in **the Office** at least some of the premises are evaluated in a different contexts to the conclusion. If Agglomeration is supposed to be *valid*, then, in some context, at least one of the premises must be *false*. So at least one of the *negations* of the premises in **the Office** should have a true reading.

But no true reading immediately attaches to any of the sentences:

(19) It's not true that Alice should be in work today.

(20) It's not true that Bob should be in work today.

(21) It's not true that Carol should be in work today.

...

(22) It's not true that Zadie should be in today.

Nor can it be elicited with further pressure. A sentence like

(23) It's not true that *Alice* should be in work today. (After all, not everyone will be in.)

still sounds false. Notice here the contrast with **Chariots**: there we can easily hear a true reading of a sentence like

(24) It's not true that *Attila* ought to swerve. (After all, *Genghis* could swerve.)

What if the missing false readings are elusive in the sense of Lewis (1996)? Whenever I consider an individual premise or its negation, this would move me into a context where it is true, and where some other premise I am not considering is false. The false readings would exist, but never where we are looking.[5]

However, if Agglomeration were valid, then no matter the context, there would be *some* person of whom one of the premises is false. So the following wide-scope *universal* claims should be univocally false:

(25) Everyone is such that they should be in work today.

(26) Everyone is someone who should be in work today.

After all, in every context, the universal quantifier has an (elusive) counterexample. Put another way, the elusive strategy says it should be univocally *true* to say:

(27) Not everyone is someone who should be in work today.

In every context there should be (elusive) counterexamples. But these are all the

[5]Hawthorne (2002) explores this response to the preface for the case of 'knows'.

wrong judgements: (25) and (26) are in fact true; (27) is false.

3 The Predicament

Agglomeration seems to fail for epistemics, but not for deontics. Both the classic view and contrastivist views fail to explain this difference.

On all versions of the classic view, 'ought' is a universal quantifier over worlds. To fix ideas, let's take von Fintel (2012)'s Kratzerian version of this view:

(28) $[\![\text{ought } \phi]\!]^{c,w} = 1$ iff $\forall w' \in BEST(w)$: $[\![\phi]\!]^{c,w'} = 1$

That is, ⌜ought ϕ⌝ is true iff ϕ is true throughout some set of best worlds. What makes a world best does not matter: (28) alone dictates the logic of 'ought'.

If the classic semantics is meant to give a unified semantics for 'ought', then Agglomeration should be valid generally. It is easy to see why: if ⌜ought ϕ⌝ and ⌜ought ψ⌝ are true, then all the best worlds are ones where ϕ is true and where ψ is true. But then all the best worlds must be ones where ⌜$\phi \wedge \psi$⌝ is true. So ⌜ought $(\phi \wedge \psi)$⌝ is true. This reasoning assumes nothing about the flavour of the 'ought'. So *any* counterexample to Agglomeration undermines the classic view, when understood as a unified semantics.

The classic semantics has limited options. One could deny that we should strive for a unified semantics for 'ought'. But, as we noted already, an ambiguity strategy does not look promising. Epistemic 'ought's occur in a number of languages, something an ambiguity strategy does not predict. Moreover, a special ambiguity treatment for 'ought' sits badly with the known fact that modal vocabulary across languages tends to express a variety of different flavours.

One could instead deny there really is a difference in the logic of epistemic and deontic 'ought's. Even putting aside my earlier responses, this misses something important. Counterexamples to epistemic Agglomeration are extremely natural and easy to find: we noted they have the structure of the preface paradox. But even *prima facie* counterexamples to deontic Agglomeration are rare and judgements about them fragile. The ease with which we can find counterexamples in the one case but not in the other strongly suggests a genuine divergence between epistemic and deontic ‘ought’s.

Contrastivist accounts of ‘ought’ fare no better. Contrastivists think that ⌜ought ϕ⌝ is true iff ϕ is *better* than its alternatives.[6] Slightly more precisely, where $ALT(\phi)$ is the set of ϕ’s alternatives and $\prec_{w,f,g}$ some ordering over propositions:

(29) $[\![\text{ought } \phi]\!]^{c,w,f,g} = 1$ iff for every $\psi \in ALT(\phi)$: $\phi \prec_{w,f,g} \psi$

There are different views on what a proposition’s alternatives are. But typically closely related propositions are allowed to have different alternatives. In particular, the alternatives for ϕ and for ψ can be disjoint from those for $\phi \wedge \psi$.

Many contrastivists share my aim of invalidating Agglomeration.[7] They predict this when ϕ is better than its alternatives, ψ is better than its alternatives but $\phi \wedge \psi$ is not better than its (distinct) alternatives. The trouble is that contrastivist views tend violate *deontic* Agglomeration. On most natural ways of thinking about alternatives, the ϕ and ψ can both be *deontically* better than their alternatives, while $\phi \wedge \psi$ is not.

For example, Jackson says that $ALT(\phi)$ is just $\neg\phi$; and that ϕ is better than ψ if the closest world where ϕ is true is better than that where ψ is true.[8] This makes **Chariots** a counterexample to Agglomeration. The premises of the inference are true: in the closest world where Genghis swerves Attila remains on course and this is better than the closest world where Genghis doesn’t swerve and so both remain on course; likewise for the closest worlds where Attila swerves and where he does not. But the conclusion is false: neither swerving leads to the worst possible outcome and so the closest world where both swerve is worse than that where at least one does not.

Of course this is all by design. **Chariots** is meant to motivate Jackson’s view. But this is not what the data support, I argued. Only the epistemic case gives robust counterexamples to Agglomeration.

[6] In addition to Jackson (1985), see also Goble (1996), Finlay (2009), Lassiter (2011) and Snedegar (2012).

[7] Many, but not all. Cariani (2013a), Cariani (2013b), and Cariani (2016b) defend a contrastivist semantics which validates Agglomeration in full generality.

[8] See also Cariani (2016a), which gives a clear account of why, on various natural ways of thinking about alternatives, Agglomeration fails on the contrastivist accounts of Finlay (2009) and Lassiter (2011), among others.

4 The View

Now let's move to consider my view. Since there are a number of moving parts, I will proceed in three steps. In this section, I will introduce my basic semantic entry for 'ought' and show how it allows Agglomeration to fail. Then I add some new elements to help distinguish between the epistemic and deontic 'ought's. The next section considers the pragmatics of these new elements.

4.1 The Semantics

On my theory, ⌜ought ϕ⌝ says that ϕ is entailed by *some best proposition*. Let's state that more precisely. Like in the standard Kratzerian semantics, we add a modal base, f, to the index. This is a function from worlds to sets of worlds, representing the background information against which we evaluate an 'ought'. We add another function g to the index also, an ordering function. This takes a world and a modal base and return an ordering over propositions, $\precsim_{w,f,g}$.

We can then define the set of best propositions that entail the information in the modal base at a world:

(30) $PBEST(w, f, g) = \{p \subseteq f(w) : p \neq \varnothing \text{ and } \neg\exists q \subseteq f(w) : q \prec_{w,f,g} p\}$

This allows us to state the semantics precisely:

(31) $[\![\text{ought } \phi]\!]^{c,w,f,g} = 1 \text{ iff } \exists p \in PBEST(w, f, g) : \forall w' \in p : [\![\phi]\!]^{c,w',f,g} = 1$

⌜ought ϕ⌝ is true just in case there is *some* proposition at the top of the ordering that entails ϕ.

What kinds of orderings does the semantics use? There is a large and, to my mind, convincing literature that a semantics for 'ought' should aim for deontic neutrality. Best just to say various different sources of value can supply an ordering to the semantics when the context is right.[9] Our judgements of betterness in context will tell us how propositions are ordered deontically.

I will be neutral about epistemic orderings too, but there are fewer plausible candidates. Epistemic orderings could track whether or not the *probability* of a

[9] See, among others, Carr (2015), Charlow (2016) and Cariani (2016b).

proposition passes a contextually supplied threshold.[10] Here is a way of formalising this:

> $\phi \precsim_{w,f,g} \psi$ iff, where θ is a contextually determined threshold, one of the following conditions holds:
>
> 1. $P(\phi | \bigcap f(w)) > \theta$; or
> 2. $P(\phi | \bigcap f(w)) \geq P(\psi | f(w))$

No proposition will be strictly better than ϕ whenever the probability of ϕ passes the given threshold, allowing ⌜ought ϕ⌝ to be true, even when we are not certain that ϕ. Another thought, touched on earlier, is that the epistemic 'ought' uses a *normality* ordering. This is more straightforward: we would have $\phi \precsim_{w,f,g} \psi$ just in case ϕ is at least as normal as ψ.

It is easy to see how this account permits Agglomeration failures. Take an ordering where ϕ and ψ are the only best propositions, so that $\phi \wedge \psi$ is further down the ranking. In such a case ⌜ought ϕ⌝ and ⌜ought ψ⌝ will be true; ⌜ought $(\phi \wedge \psi)$⌝ will not.[11]

[10]Because of the arguments in Yalcin (2016), this can at best be a necessary condition for an epistemic 'ought'. A plausible additional constraint is that learning the negation of the prejacent would also lower one's credence in some salient background proposition. I ignore these complications here.

[11]Notice that, as stated, my view validates the controversial inference pattern known as **Inheritance**:

Inheritance. If $\phi \vDash \psi$ then ⌜ought ϕ⌝ $\vDash$ ⌜ought ψ⌝

This inference is often rejected on the basis of examples from Ross (1941):

(i) a. I ought to mail the letter.
b. $\not\rightsquigarrow$ I ought to mail or burn the letter.

I am convinced by von Fintel (2012) that Inheritance is actually desirable and that Ross's examples should be explained pragmatically. But, whether or not this is right, this is somewhat orthogonal to the discussion here. We could restate the view so that ⌜ought ϕ⌝ is true, not when ϕ is entailed by a best proposition, but rather when ϕ *is* a best proposition. Or we could instead adopt a non-Boolean disjunction, as, for instance, Alonso-Ovalle (2009), Fusco (2015) and Goldstein (2019) do to handle this or related problems. Either would be sufficient to invalidate Inheritance.

4.2 Restrictions on the Ordering

To secure deontic but not epistemic Agglomeration, we need tools to help select the right orderings in each case. I add two new elements to the semantics: a question parameter that restricts the propositions quantified over; and a consistency constraint.

4.2.1 Questions

The set of propositions is very large indeed. Which ones get ordered in our semantics? All of them? Or merely some of them?

Throughout a conversation, certain distinctions matter and others do not. I suggest the ordering should not be sensitive to distinctions we are not attending to in context. If a proposition makes distinctions we are not interested in, there is no point in ranking it.

We can capture this in the semantics by appeal to what I call a *relevance question*.[12,13] By choosing a question whose complete answers do not distinguish between certain worlds, we can model when propositions are relevant. A proposition is relevant if it is a complete answer to the question or at least a union of complete answers. Otherwise it crosscuts or makes finer distinctions than the distinctions we are making in context and so is not relevant. I assume that it is this kind of question that gets supplied to the semantics.

The semantics of questions helps us cash this out. On a natural semantics for questions, they themselves are sets of propositions: a question partitions the set of worlds into the propositions giving complete answers to the question.[14] For instance the question *Who out of Alice and Billy had ice cream?* is the set

{*Just Alice had ice cream*, *Just Billy had ice cream*, *Both had ice cream*,

[12]In [omitted] I put a similar notion to work in determining orderings for counterfactuals.

[13]I take the relevance question to be different from Roberts (1996)'s *question under discussion* (though is probably ultimately related to it). For suppose we want to answer the question whether Billy came to the party at all last night. Now compare the propositions *Alice came to the party* and *Alice came to the party for an hour*. Neither proposition is an answer to the question under discussion. But there is still a felt difference between the two: the second introduces distinctions that are in some sense not relevant at all in the context; the second does not. The relevance question functions more like a standard of precision parameter than a question under discussion.

[14]Here I use Karttunen (1977)'s account of questions to model an issue relevant in the context. I take no stand on the semantics for interrogatives.

Neither had ice cream$\}$.

It will also need the notion of *partial answer* to Q. Say that p is a partial answer to Q iff p is a union of some elements of Q. So, for instance, among the partial answers to the question *Who out of Alice and Billy had ice cream?* will be the proposition *either just Alice or both Alice and Billy had ice cream*.

These relevant questions supply us with the set of propositions that get ordered. Where Q is the relevant question, g orders just the partial answers to Q. This gives us a fairly sensible answer to our opening question: the final ordering only cares about propositions distinguished by the relevant question. A proposition gets ordered so long as it does not cross-cut distinctions made by our relevance question.

We will represent this formally by defining a new kind of ordering function. First, we let ordering functions also take *questions* as arguments: g is now a function from worlds, modal bases and *questions* to an ordering $\prec_{w,f,g,Q}$. We can then carve out the set of *question-sensitive* ordering functions as follows:

> **Question-sensitivity:** g is a *question-sensitive* ordering function iff $\prec_{w,f,g,Q}$ orders only complete and partial answers to Q.

So, rather than drawing from the set of ordering functions more generally, our semantics will draw from the set of question-sensitive ordering functions.

Let's redefine $PBEST$ with this in mind. Say that $Q|S$ is the partition imposed by Q restricted to worlds in the set S; that is,

$$Q|S = \{p : \exists q \in Q \text{ and } p = q \cap S\}.$$

Then the set of best propositions as follows:

(32) $$PBEST(w,f,g,Q) = \{p \subseteq f(w) : p \neq \varnothing \text{ and } \neg\exists q : q \prec_{w,f,g,Q|f(w)} p\}$$

Crucially, question-sensitivity will allow us to give a principled account of what propositions get ordered and to state intuitive constraints on orderings.

4.2.2 Definedness Conditions

I propose that, whatever the relevant ordering is, 'ought' requires the best propositions to be pairwise *contextually consistent*. I adopt this as a definedness constraint: given a particular ordering, ⌜ought ϕ⌝ has a truth-value only if every pair of best propositions are consistent, given the background information. I spell this out formally as follows:

> **Consistency:** $[\![\text{ought } \phi]\!]^{c,w,f,g,Q}$ is defined only if when ψ and ρ are in $PBEST(w, f, g, Q)$, then $\psi \wedge \rho$ is consistent with $f(w)$.

(Note that this constraint does not require that $\psi \wedge \rho$ actually be a member of $PBEST(w, f, g)$.)

It is important that we use *pairwise*, rather than overall, consistency. Returning to **the Office** illustrates why. There, on the natural way of construing the situation, the entire set of best propositions is

{*Alice is in work today*, *Billy is in work today*, ..., *Somebody is absent*}

This set is pairwise consistent because any two people could be in together; and any particular person could be in, even if somebody is absent. But it is not overall contextually consistent because, given the scenario, somebody is out of work just in case one of Alice, Billy, Carol, and so on, is absent.

What kind of definedness or undefinedness is relevant here? I will think of it along the lines of a definedness condition on a pronoun.[15] Pronouns can carry gender, person or number features. Following Kratzer and Heim (1998), semanticists tend to model these as definedness constraints. For instance, the extension of a pronoun like 'she' will only be defined on a given variable assignment if the extension is female. Since it is commonly thought that, in the final analysis, items like modal bases and ordering functions will also supplied by variable assignments, this kind of definedness seems appropriate for constraining orderings.[16]

Why should 'ought' include such a constraint?[17] Well, in the deontic case, the things we ought to do should be consistent with each other: after all, 'ought's are

[15] Here I follow Mandelkern (2019)'s gloss on his locality constraint for epistemic modals.

[16] Though, for simplicity, I do not formulate the view using variable assignments.

[17] Are there any precedents for this particular definedness constraint? I do not know of any, but I

supposed to be action guiding. But this cannot be simply built into the meaning of 'ought': as we saw, we do not want such a restriction for epistemics. I suggest that the pairwise constraint is a way to generate this restriction without interfering with epistemics: as we will see, it effectively gets strengthened to an overall consistency constraint just for deontics.

4.3 Summing up

Let's put this all together. The points in our model are tuples of worlds, modal bases, question-sensitive ordering functions and questions. Then we have:

(32) $PBEST(w, f, g, Q) = \{p \subseteq f(w) : p \neq \varnothing \text{ and } \neg\exists q \subseteq f(w) : q \prec_{w,f,g,Q|f(w)} p\}$

(33) a. $[\![\text{ought } \phi]\!]^{c,w,f,g,Q}$ is defined only if when ϕ and ψ are in $PBEST(w, f, g, Q)$, then $\phi \wedge \psi$ is consistent with $f(w)$.

b. If defined, $[\![\text{ought } \phi]\!]^{c,w,f,g,Q} = 1$ iff $\exists p \in PBEST(w, f, g, Q) : \forall w' \in p : [\![\phi]\!]^{c,w',f,g,Q} = 1$

5 Pragmatics

Before getting to the predictions, I outline and motivate some important claims about the pragmatics of the relevance question and orderings.

5.1 Questions and Orderings

The apparatus of questions allows us to state a fundamental difference between epistemic and deontic orderings.

can think of at least one other place where it would be useful. As Yalcin (2010) notes, sometimes we judge outcomes probable when they are more probable than some threshold below 0.5, one fixed by the relevant alternatives. This is known as the alternative outcomes effect. As Yalcin suggests, a natural way to model this makes 'probably' a gradable adjective and allows the threshold to sometimes fall below 0.5. However, without further constraints this allows for both ⌜probably ϕ⌝ and ⌜probably $\neg\phi$⌝ to be true. A natural fix here is to say that ⌜probably ϕ⌝ is true if one of a contextually supplied set of propositions passes the threshold and entails ϕ; and to add that, however those alternatives are picked, the set of most probable alternatives must obey something like my consistency constraint.

Fix on a background question Q. Now take some proposition ϕ which is the disjunction some cells $\pi_1, ..., \pi_n$ of the Q partition. How does ϕ compare to its constituent Q-cells? Must some constituent cell be at least as good as ϕ? Or can they all be *worse* than ϕ?

The answer differs importantly depending on whether the ordering is epistemic or deontic. Take the epistemic first. It is natural to think that a proposition can be epistemically worse than its constituent cells. For example, if epistemics are based on probability, then ϕ can clearly be ranked higher than any of its constituent Q-cells: a disjunction will usually be more probable than either of its disjuncts.

With deontics it is different: ϕ cannot be better than *all* of its constituent Q-cells. We can think of $\pi_1, ..., \pi_n$ as the ways that ϕ can come about. Now ϕ may well be better overall than many or even most of its realisations. But ϕ cannot be better than *all* of its possible realisations: at least one Q-cell must be at least as good as ϕ. Likewise, it cannot be *worse* than all ways for it to come about: its value must lie somewhere between the best and worst ways for it to come about.

Summing this all up, we have:

> If g is deontic, then if Π is a partial answer to Q, then, where $\kappa, \kappa' \subseteq \Pi$ are respectively best and worst complete answers that entail Π, $\kappa \precsim_{w,f,g,Q} \Pi \precsim_{w,f,g,Q} \kappa'$.
>
> If g is epistemic and Π is a partial answer to Q, then there may be no complete answer $\kappa \subseteq \Pi$ such that $\kappa \precsim_{w,f,g,Q} \Pi$.

This difference in the properties of deontic and epistemic orderings is exactly why Agglomeration will fail only in the epistemic case.

5.2 The relevant question

We have said that the relevance question dictates the fineness of the ordering. But which *particular* questions get to be the relevance question?

In the epistemic case I take it to be the main question which the speakers are aiming to answer. Whatever this question is, it will be at least as fine-grained as the polar question corresponding to the prejacent of the 'ought' claim. That is, for a sentence like,

(34) Sarah ought to be home any minute.

the question is at least as fine-grained as the question

$\{$*Sarah will be home any minute*, $\neg$*(Sarah will be home any minute)*$\}$

The deontic case is somewhat more complex. Here I think there are *two* kinds of relevance questions.

One makes at least as many distinctions as there are complete outcomes that we are interested in. Take for instance the question *what did/will happen?*. This question partitions the modal base into propositions that state what happens in the relevant scenario. Take, for example, **Dessert**. Here the question *what will happen?* is the partition

> $\{$*I have just pie*, *I have just cannoli*, *I have just cheesecake*, *I have both pie and cannoli*, *I have both pie and cheesecake*, *I have both cannoli and cheesecake*, *I have all three*, *I have none*$\}$.

Each total outcome is represented in its own proposition.

The other kind of question simply asks *how good was/will the outcome be?* and does not distinguish between equally good outcomes. This question partitions the modal base into propositions which say that the agent got one of the best outcomes, one of the second best outcomes, and so on. In general, the partition will look something like this:

> $\{$ *I get a best option*, *I get a second best option*, ..., *I get a worst option*$\}$

Now, remember are partitioning modal bases, and not the whole space of worlds. So, for instance, in **Dessert**, the partition will amount to something like:

> $\{$*I have just pie or just cannoli*, *I have just cheesecake*, *I have pie and cannoli or no dessert at all*, *I have pie and cheesecake*, *I have cannoli and cheesecake*, *I have all three*$\}$

The different kinds of question reflect different ways we can use outcomes to distinguish between possibilities. *What will happen?* distinguishes between them exactly as finely as our options do. *How good will the outcome be?* only distinguishes between possibilities up to how good the outcomes are, lumping possibilities with equally good outcomes into one cell. As a result, *What will happen?* tends to partition more finely than *How good will the outcome be?*. This will help predict Indifference by filtering out of an ordering propositions that violate Agglomeration.[18]

6 Predictions

Now we are in a position to explain our two pieces of data:

Aggomeration failure. Counterexamples to Agglomeration arise only for epistemic 'ought's.

Indifference. In cases where ϕ_1, ... ,ϕ_n are jointly (intuitively) best, $\ulcorner$ought $\phi_1 \vee ... \vee \phi_n\urcorner$ is true.

I first show in the abstract why they are predicted. Then I work through the particular kinds of cases we have seen.

6.1 Predictions in the Abstract

Agglomeration fails only for epistemics because of how epistemic and deontic orderings interact with the consistency constraint. Best deontic propositions drag at least one complete answer up to their level but epistemics need not. This ensures

[18] Is there any independent motivation for using both fine-grained and coarse-grained questions in the semantics for deontic vocabulary? I think we may find some in how we talk about *options* more generally. The following sound interchangeable:

(i) Going right or left is the best option.

(ii) Going right and going left are the best options.

However, this is puzzling given that singular 'the' usually presupposes or entails uniqueness. We can explain this by connecting relevance questions to how we discriminate between options: the sense of option in (i) is determined by a coarse-grained question; that in (ii) is determined by a fine-grained question.

Agglomeration failures in the deontic case yield a (pairwise) inconsistent set of best propositions; while epistemic Agglomeration can fail without inconsistency.

Here is a proof sketch.[19] For contradiction, assume that we have a case of deontic Agglomeration-failure: ⌜ought ϕ⌝ and ⌜ought ψ⌝ are true, but ⌜ought $(\phi \wedge \psi)$⌝ is false. By my semantics, there must be some Π_1 and $\Pi_2 \in PBEST(w, f, g)$ such that Π_1 entails ϕ and Π_2 entails ψ. Now since ⌜ought ϕ⌝ and ⌜ought ψ⌝ are both true, the consistency constraint must be satisfied: the intersection of Π_1 and Π_2 must be compatible with the background information.

Since ⌜ought $(\phi \wedge \psi)$⌝ is to be false, $\Pi_1 \cap \Pi_2$ cannot be ranked at least as high as Π_1 or Π_2. It is at this step the property of deontic orderings becomes important. There must be some other $\pi_1 \subseteq \Pi_1$ and $\pi_2 \subseteq \Pi_2$ that are at least as good as Π_1 and Π_2 respectively. π_1 and π_2 cannot be be subsets of$\pi_2 \subseteq \Pi_2$, for we are supposing Agglomeration fails here. But then π_1 and π_2 cannot be consistent given the background information: π_1 must come from the $\neg(\Pi_1 \cap \Pi_2)$ region of Π_1 and likewise for π_2 and Π_2. So the consistency constraint is not met after all and ⌜ought ϕ⌝ and ⌜ought ψ⌝ are not true. Contradiction.

Now let's see why Agglomeration can fail for epistemics. Suppose ⌜ought ϕ⌝ and ⌜ought ψ⌝ are true. Then there must be some Π_1 and $\Pi_2 \in PBEST(w, f, g)$ such that Π_1 entails ϕ and Π_2 entails ψ; and Π_1 and Π_2 must satisfy the consistency constraint. Again, we must suppose that $\Pi_1 \cap \Pi_2$ is not among the best propositions. But because the ordering is epistemic, this does not mean that some *other* cell must be at least as good: in fact, every cell composing Π_1 could be worse than it, and likewise for Π_2. So the consistency constraint can be satisfied in this situation: though Π_1 and Π_2 are best, none of their constituent cells need be.

Why do we predict Indifference? When we have multiple best options ϕ_1, ..., ϕ_n, there must be best complete answers $\pi_1, ..., \pi_n$ that entail each. Since these answers will be pairwise inconsistent, a question which includes both will yield undefinedness. I suggest the pragmatics will search for a question which does not yield undefinedness. We have seen one such question, the question *how good is the outcome?*: this question will not rank ϕ_1, ..., ϕ_n but only their disjunction. By our deontic constraint, that disjunction, $\phi_1 \vee ... \vee \phi_n$, must also be among the best

[19]The full proof is provided in the Appendix.

propositions: it must fall between the best and world complete answers that entail it, but all are best. Since the rest of the ordering remains unchanged, $\phi_1 \vee ... \vee \phi_n$ is among the best propositions and $\ulcorner$ought $\phi_1 \vee ... \vee \phi_n\urcorner$ is true.

6.2 Case by case

Now let's work through the cases.

Start with **the Office**. We know that each person comes into work far more often than not. Regardless of whether the epistemic ordering is probability- or normality-based, plausibly the set of best propositions looks as follows:

> {*Alice is in work today*, *Billy is in work today*, ... , *Zadie is in work today*, *Not everybody is in work today*}

These propositions are pairwise contextually-consistent: any pair could be in work together; and it is consistent that both a given person is in work but not everybody else is. Importantly, it's possible for this set to be the *entire* set of best propositions: epistemic orderings do not require us to place any finer-grained propositions among the best.

This set of best propositions correctly predicts the truth of

(3) Alice should be in work today.

(4) Bob should be in work today.

(5) Carol should be in work today.

...

(6) Zadie should be in today.

and the falsity of

(7) Everyone should be in today.

There is a best proposition that entails Alice will be in work, that entails Bob will be work, and so on. But no best proposition entails everybody will be in work.

Now recall **Going Home**:

> **Going Home**. There are three different routes home, A, B and C. A and B get us home equally quickly, but C takes longer. All we care about is how quickly we get home.

Avoiding Agglomeration failure is straightforward here. If the relevance question is *What will happen?*, the propositions *I take route A* and *I take route B* will be ranked at the top. But clearly such an ordering does not satisfy the consistency constraint: I can't take both routes at once. So neither of the sentences

(35) I ought to take route A.

(36) I ought to take route B.

will be true.

When instead the relevance question is *How good will the outcome be?*, we predict **Indifference** holds here. That question imposes the following partition on the modal base:

> {*I take route A or B*, *I take route C*}

Plainly, *I take route A or B* will be among the best partial answers based on this partition. This is enough to predict the truth of the disjunctive 'ought',

(37) I ought to take route A or B.

Moreover, it is reasonable to think that this is the *only* reading speakers should get here. When faced with one ordering that yields definedness and another that does not, speakers and audiences should select the one that does.

Finally, let's turn to **Dessert**. As we saw, the relevance question *What will the outcome be?* gives us the partition:

> {*I have just pie*, *I have just cannoli*, *I have just cheesecake*, *I have both pie and cannoli*, *I have both pie and cheesecake*, *I have both cannoli and cheesecake*, *I have all three*, *I have none*}.

Again, because of the set-up of the case, the ordering of partial answers based on this partition will be inconsistent. Presumably *I have pie* and *I have cannoli* will be among the best propositions based on this ordering. But then *I have only pie* and *I have only cannoli* among the best propositions too, given that *I have both pie and cannoli* certainly is not. This yields undefinedness.

As before, we predict that **Indifference** holds through accommodation. The relevance question *How good will the outcome be?* induces something like the following partition on the modal base:

> $\{$*I have just pie or just cannoli*, *I have just cheesecake*, *I have pie and cannoli or no dessert at all*, *I have pie and cheesecake*, *I have cannoli and cheesecake*, *I have all three*$\}$

Here the proposition *I have pie or cannoli* will be among the best propositions. Again, charity forces us to choose the ordering that yields definedness and so we predict the truth of

(38) I ought to have pie or cannoli.

Working through the cases confirms what we saw in the abstract. Agglomeration can fail for epistemics: they bring no violation of consistency. It cannot fail for deontics: there a violation of consistency would be required. Instead, context selects the alternative ordering that satisfies 'ought's definedness conditions; and, in cases with multiple best options, that ordering will deliver Indifference.

7 Dilemmas

Some think moral dilemmas are genuine counterexamples to deontic Agglomeration. What should my theory say about them?

Take the following case:[20]

> **Sophie's Choice.** Sophie is forced to choose which of her children is going to be sent to the labour camp and which is going to be killed. If she chooses neither child, then both will be killed.

[20] Sinnott-Armstrong (1985) calls this a *symmetric dilemma*, as there is no morally relevant difference between the sources of the obligations (unlike in the case of Sartre's famous student).

Here it is possible to get a true reading of both of the following sentences.

(39) Sophie ought to save her daughter.

(40) Sophie ought to save her son.

But obviously the agglomerated 'ought'

(41) Sophie ought to save her daughter and her son.

is not true here, for Sophie cannot do both of these things together. So, if all of these are evaluated in the same context, we have a failure of Agglomeration. I do not predict this.

I will explore two reactions: one is to generalise my framework; another is to use contextualism to handle dilemmas separately from the kinds of Agglomeration failures I have discussed. I think both are defensible, but favour the latter.

7.1 Modifying the Consistency Constraint

One option is to weaken the consistency constraint by appeal to incomparability.

As von Fintel (2012) notes, dilemmas do not simply arise whenever we have more than one best option; cases like **Going Home** or **Dessert** are clearly not dilemmas. Following Fraassen (1973), Horty (2012) and Swanson (2016), I suggest that what is distinctive about dilemmas is that the best options are *incomparable*.[21]

Why think this? Because *mild sweetening* of an option does not resolve the dilemma.[22] Suppose Sophie can choose between having her daughter saved and having her son saved *and* treated slightly better in the labour camp. Having her son saved *and* treated slightly better in the labour camp is better than just having her son saved; but improving this option does not resolve the dilemma. So saving her son and saving her daughter cannot be equally good. Neither option is *better* than the other, so we should conclude the two options are incomparable.

[21] I mean this simply in the formal sense that, for the contextually supplied ordering, neither is at least as good as the other. Some ethicists, like Chang (2002), make further distinctions among kinds of incomparability.

[22] The name is Hare (2010)'s.

In light of this, we could weaken the consistency constraint to say merely each pair of *comparable* best propositions that must be consistent. More formally:

> **Comparable Consistency:** $[\![\text{ought } \phi]\!]^{c,w,f,g,Q}$ is defined only if for every *comparable* ψ and ρ in $PBEST(w, f, g)$, $\psi \wedge \rho$ is consistent with $\bigcap f(w)$.

In **Sophie's Choice**, the best propositions are *Sophie saves her daughter* and *Sophie saves her son*. Though they are contextually inconsistent, the new constraint predicts that (39) and (40) are true. *Sophie saves her daughter* and *Sophie saves her son* are not comparable; so this ordering can be supplied to the semantics without undefinedness.

7.2 Rejecting Dilemmas

We saw how to tweak framework my framework, if dilemmas are genuine Agglomeration failures. However, I am sceptical that they are.

von Fintel (2012) observed that apparent dilemmas arise for 'must' as well as 'should'. In **Sophie's Choice** we just as easily could have considered:

(42) Sophie must save her daughter.

(43) Sophie must save her son.

and noted that both seem to have true readings here.[23] To really account for dilemmas, our account must extend to 'must' as well.

But there is good reason to think that these apparent dilemmas for 'must' are not genuine Agglomeration failures. As Sinnott-Armstrong (1985) and Horty (2012) note, dilemmas yield contradictions assuming natural principles for deontic 'must's and 'may's:

Duality: $\ulcorner$must $\phi\urcorner \vDash \ulcorner\neg$ may $\neg\phi\urcorner$ and $\ulcorner\neg$ may $\neg\phi\urcorner \vDash \ulcorner$must $\phi\urcorner$

Must-to-may: $\ulcorner$must $\phi\urcorner \vDash \ulcorner$may $\phi\urcorner$.

Must-Inheritance: If $\phi \vDash \psi$, $\ulcorner$must $\phi\urcorner \vDash \ulcorner$must $\psi\urcorner$

[23] Some hear dilemmas for strong necessity modals more easily by using 'have to' instead of 'must'.

If we have a genuine dilemma, then, for some ϕ and ψ where $\phi \vDash \neg\psi$, ⌜must ϕ and must ψ⌝ is true (all in the one context). Assuming the above principles, this entails ⌜may $\neg\phi$ and $\neg$(may $\neg\phi$)⌝ is true. This of course cannot be. Of all the assumptions above, the existence of dilemmas looks to me to be the weakest.[24]

Context-sensitivity is the obvious tool to reach for. If (42) is evaluated in a different context, and so against a different ordering, to (43), then both can be true without denying any of the principles above and without contradiction. Naturally, this explanation extends to 'ought' too, relieving the pressure on my consistency constraint. That constraint only applies within, not across, contexts.

There is also an independent reason to favour a context-shifting explanation. 'Must' is what I will call *pseudo-factive*: as Ninan (2005) observes, it sounds incoherent to say

(44) #You must clean your room, even though you aren't going to.

Though they surely don't entail them, 'must's seem to carry some commitment to the truth of their prejacent. This makes the existence of dilemmas for 'must' somewhat surprising. If *both* (42) and (43) are true, then pseudo-factivity says there should be a felt commitment to both Sophie saving her daughter and saving her son. But even though at most one of these can hold, both (42) and (43) are acceptable.

I want to suggest this is further evidence for a multiple orderings approach. Notice that when there are multiple orderings in play, the pseudo-factivity of 'must' disappears. If I say something like:

(45) According to your father, you must be home by 9; and according to your mother you must be in bed by 10

it does not seem to carry the same commitment to the prejacent. To see this observe that I can continue with something like:

(46) But of course, you're not going to do either of those things.

When there are multiple sources of obligation in play and we have not committed ourselves to any one of them, 'must' does not seem to carry the same commitment

[24]Note that only a fairly trivial instance of Must-Inheritance is used here.

to the prejacent.

Now if both (42) and (43) were interpreted relative to the same ordering we should expect felt commitments to two contradictory propositions. On the other hand, if there were multiple orderings in play, neither of which were completely endorsed or taken to be binding, we would predict such commitments. Since the latter is what we actually see, the multiple ordering strategy does well here.

In sum, while my framework can accommodate dilemmas, I think for independent reasons they should not be seen as genuine Agglomeration failures.

7.3 Interlude: Comparison to Previous Treatments of Dilemmas

While we are on the subject of dilemmas, it is worth asking how my account differs from the standard semantic treatment in dilemmas, elaborated in Fraassen (1973), von Fintel (2012), Horty (2012) and Swanson (2016). Following Horty (2012), call this the *conflict account*.

Take von Fintel (2012)'s Kratzerian presentation of the view. We have an ordering source g that collects the set of prima facie 'ought's at w. For example, in **Sophie's Choice**, $g(w)$ would be $\{$*Sophie saves her daugher*, *Sopie saves her son*$\}$. Say that $D(g(w), f(w))$ is the set of maximal contextually-consistent subsets of $g(w)$.[25] Then 'ought' has the following entry:

(47) $[\![\text{ought } \phi]\!]^{c,w,f,g} = 1$ iff there is some $S \in D(g(w), f(w))$ such that for all $w' \in \bigcap S$: $[\![\phi]\!]^{c,w,f,g} = 1$.

This looks very close indeed to my view, especially in its existential quantification. But mine is less restrictive, quantifying simply over the best propositions in some ordering. This is to my advantage. The conflict account does not generalise to epistemic Agglomeration failure. Dilemmas involve *pairwise* inconsistency in what ought to be; epistemic Agglomeration failures involve mere *overall* inconsistency in what ought to be. And the conflict account is specifically tailored for dilemmas.

The most natural way to apply the conflict account to cases like **the Office** is for $g(w)$ to contain the propositions which we think are sufficiently likely or normal.

[25]In other words, the set of subsets S such that $S \cap \bigcap f(w)$ is consistent and there is no S' such that $S \subset s' \subseteq g(w)$ and $S' \cap \bigcap f(w)$ is consistent.

Naturally that set will be inconsistent. Its maximal consistent subsets will be:

$\{$*Billy is in, Carol is in, Daniel is in, ..., Zadie is in, not everyone is in*$\}$

$\{$*Alice is in, Carol is in, Daniel is in, ..., Zadie is in, not everyone is in*$\}$

$\{$*Alice is in, Billy is in, Daniel is in, ..., Zadie is in, not everyone is in*$\}$

$\{$*Alice is in, Billy is in, Carol is in, Daniel is in, ..., Zadie is in*$\}$

The conflict account predicts conclusion of the Agglomeration inference has a true reading here. Our premises from before,

(3) Alice should be in work today.

(4) Bob should be in work today.

(5) Carol should be in work today.

...

(6) Zadie should be in today.

are all true here. But, because of the last maximal consistent set on our list above, the conclusion is also true:

(7) Everyone should be in today.

There is after all a maximal consistent set which entails that everybody is in today, the one that leaves out *not everybody will be in today*, and includes the rest of $g(w)$. This is not what we want: the conclusion is univocally false.

The conflict account also predicts that for each premise like (3), we should also predict a true reading of a sentence like

(48) Alice should *not* be in work today.

After all, for each worker there is a maximal consistent set that (contextually entails) they are in today. Again, this is not what we want: no such sentences are true, only the original premises.

Despite their similarity, my account and the conflict account make very different predictions in **the Office**. My semantics accounts for it with ease; the conflict account does not. This is because my account, and not the conflict account, is tailored to a more general kind of Agglomeration failure.

8 Conclusion

Neither the classic semantics nor the contrastivists get Agglomeration right. I gave a new semantics where 'ought' is an existential quantifier over best propositions and so ⌜ought ϕ⌝ and ⌜ought ψ⌝ can be true, while ⌜ought $(\phi \wedge \psi)$⌝ is false. To capture the differing entailments of deontics and epistemics, I added a layer of question-sensitivity to 'ought', so that the best propositions must also be partial answers to a background question; I also added a pairwise consistency constraint. Together with some assumptions about the pragmatics of the background question and orderings, these deliver deontic but not epistemic Agglomeration.

In closing, let me speculate on how my semantics fits into the more general typology of modals. Quantifying over higher types has also been proposed for *possibility* and *(strong) necessity modals*, by Moss (2015), Mandelkern et al. (2017) and Khoo (forthcoming), among others. There is a particular continuity between my view and that of Mandelkern et al. (2017). They take ability modals to quantify over *actions* saying that ⌜S can ϕ⌝ is true iff there is some (practically available) action A such that if S tried to A, S would ϕ; and ⌜S must ϕ⌝ is true iff all practically available actions A are such that if S tried to A, they would ϕ.

A natural generalisation of this idea substitutes *best propositions* for available actions, with possibility and necessity modals getting the following entries:

(49) $[\![\text{can } \phi]\!]^{c,w,f,g} = 1$ iff $\exists p \in PBEST(w,f,g) : \exists w' \in p : [\![\phi]\!]^{c,w',f,g} = 1$

(50) $[\![\text{must } \phi]\!]^{c,w,f,g} = 1$ iff $\forall p \in PBEST(w,f,g) : \forall w' \in p : [\![\phi]\!]^{c,w',f,g} = 1$

Mandelkern et al. (2017)'s semantics will result when the best propositions are the set of closest worlds where you try to do an available action; they will be the propositions that say the agent performs a representative attempt to A, relative to the world of evaluation.

When we move to higher types, there is more space for modals of intermediate strength. On my proposal, 'ought' falls exactly in this space between possibility and necessity: it is stronger than a possibility modal, because it requires entailment by a best proposition and not mere consistency; it is weaker than a necessity modal, because it does not require entailment by *all* best propositions.[26] To be sure, this kind of type-raised approach to modals is an emerging, rather than established view. But it does hold out promise to illuminate a number of independent issues, among them the inferential properties of 'ought'.

Appendix

First, we establish three important facts about questions:

Fact 1. If $\phi \cap \rho$ is a partial answer to $Q|\rho$, $\psi \cap \rho$ is a partial answer to $Q|\rho$ and $\psi \cap \rho \subset \phi \cap \rho$, then $\phi \cap \rho - \psi \cap \rho$ is a partial answer to $Q|\rho$.

Proof. Since $\phi \cap \rho$ is a partial answer to $Q|\rho$, it is the union of some set $\{\kappa_1, ..., \kappa_n\}$ of complete answers to $Q|\rho$. Since $\psi \cap \rho$ is a partial answer to $Q|\rho$ and $\psi \cap \rho \subset \phi \cap \rho$, $\psi \cap \rho$ is union of some set of complete answers which are subsets of $\phi \cap \rho$; call this set Π. Since $\Pi \subset \{\kappa_1, ..., \kappa_n\}$, $\{\kappa_1, ..., \kappa_n\} - \Pi$ is also a partial answer, being the set of complete answers in $\{\kappa_1, ..., \kappa_n\}$ but not Π. The union of this set is $\phi \cap \rho - \psi \cap \rho$.

Fact 2. For any non-empty ϕ and question Q, there is a smallest partial answer to Q, ψ, s.t $\phi \subseteq \psi$.

Proof. ψ is the union of the complete answers containing some ϕ-words. For take any partial answer ρ that is not a subset of ψ. Then there is some complete answer to Q that is a subset of ψ but not a subset of ρ. But by stipulation all and only the complete answers in ψ contain ϕ worlds. So ρ does not contain all ϕ-worlds. So $\phi \nsubseteq \rho$ after all. Contradiction.

Fact 3. If $\phi \cap \rho$ and $\psi \cap \rho$ are both partial answers to $Q|\rho$, then if $(\phi \cap \psi) \cap \rho$ is non-empty, it is a partial answer to $Q|\rho$.

Proof. Suppose for contradiction that $\phi \cap \rho$ and $\psi \cap \rho$ are both partial answers to

[26] This kind of weakness is distinct from that discussed by von Fintel and Iatridou (2008).

$Q|\rho$, but $(\phi \cap \psi) \cap \rho$ is non-empty and not a partial answer to $Q|\rho$. Consider now σ, the smallest partial answer to $Q|\rho$ that has $(\phi \cap \psi) \cap \rho$ as a subset. (We know this exists from Fact 2.) Since $(\phi \cap \psi) \cap \rho$ is not a partial answer to $Q|\rho$, there must be some complete answer $\kappa \subseteq \sigma$ such that $\kappa \cap (\phi \cap \psi) \cap \rho \neq \varnothing$ but also $\kappa \nsubseteq ((\phi \cap \psi) \cap \rho)$. Since the complete answers partition $Q|\rho$, the worlds in $\kappa \cap (\phi \cap \psi) \cap \rho \neq \varnothing$ are contained in no other complete answer to $Q|\rho$. But those worlds appear in both $\phi \cap \rho$ and $\psi \cap \rho$, which means that $\kappa \subseteq (\phi \cap \rho)$ and $\kappa \subseteq (\psi \cap \rho)$. So $\kappa \subseteq ((\phi \cap \psi) \cap \rho)$ after all. Contradiction.

Now recall the difference between deontic and epistemic orderings:

Assumption 1 (Deontic orderings). If g is deontic, then if Π is a partial answer to Q, then, where $\kappa, \kappa' \subseteq \Pi$ are respectively best and worst complete answers that entail Π, $\kappa \precsim_{w,f,g,Q} \Pi \precsim_{w,f,g,Q} \kappa'$.

Assumption 2 (Epistemic orderings). If g is epistemic and Π is a partial answer to Q, then there may be no complete answer $\kappa \subseteq \Pi$ such that $\kappa \precsim_{w,f,g,Q} \Pi$.

From all the above, we prove an important result about the semantics:

Fact 4. (No deontic Agglomeration Failure) There are no w, f, g, Q such that:

1. $[\![\text{ought } \phi]\!]^{w,f,g,Q} = 1$ and $[\![\text{ought } \psi]\!]^{w,f,g,Q} = 1$
2. g, the ordering function, is deontic;
3. and $[\![\text{ought } (\phi \wedge \psi)]\!]^{w,f,g,Q} = 0$

Proof. Suppose for contradiction that (1) and (2) are true and (3) is false.

From (1), there are $\Pi_\phi, \Pi_\psi \in PBEST(w, f, g)$ that entail ϕ and ψ respectively. Since $[\![\text{ought } \phi]\!]^{w,f,g,Q}$ is defined and true, $\precsim_{w,f,g,Q}$ satisfies the consistency constraint. So $\Pi_\phi \cap \Pi_\psi \neq \varnothing$.

From (1), $(\Pi_\phi) \cap f(w)$ and $(\Pi_\psi) \cap f(w)$ must be partial answers to Q_c, since they are ranked by $\precsim_{w,f,g,Q}$. From Fact 3 it follows that $(\Pi_\phi \cap \Pi_\psi) \cap f(w)$ is a partial answer to Q. Fact 1 then entails both $(\Pi_\phi) \cap f(w) - (\Pi_\phi \cap \Pi_\psi) \cap f(w)$ and $(\Pi_\psi) \cap f(w) - (\Pi_\phi \cap \Pi_\psi) \cap f(w)$ are also partial answers to Q

From (3)'s falsity we have that there is *no* $\Pi \in PBEST(w,f,g)$ that entails $\phi \cap \psi$. In particular, no complete answer $\kappa \subseteq (\Pi_\phi \cap \Pi_\psi) \cap f(w)$ is a member of $PBEST(w,f,g)$; every $\kappa \subseteq (\Pi_\phi \cap \Pi_\psi) \cap f(w)$ is *worse* than $(\Pi_\phi) \cap f(w)$. Similarly for $(\Pi_\phi \cap \Pi_\psi) \cap f(w)$ and $(\Pi_\psi) \cap f(w)$.

Now clearly $(\Pi_\phi) \cap f(w) = ((\Pi_\phi \cap \Pi_\psi) \cap f(w)) \cup (((\Pi_\phi) - (\Pi_\phi \cap \Pi_\psi)) \cap f(w))$. Given (2) and Assumption 1, if no $\kappa \subseteq (\Pi_\phi \cap \Pi_\psi) \cap f(w)$ is at least as good as $(\Pi_\phi) \cap f(w)$, then some $\kappa \subseteq (\Pi_\phi) \cap f(w) - (\Pi_\phi \cap \Pi_\psi) \cap f(w)$ is. Call it κ_ϕ. Similarly, some $\kappa \subseteq (\Pi_\psi) \cap f(w) - (\Pi_\phi \cap \Pi_\psi) \cap f(w)$ must be as good as $\Pi_\psi \cap f(w)$. Call it κ_ψ.

Since κ_ϕ is as good as Π_ϕ and κ_ψ is as good as Π_ψ, both κ_ϕ and κ_ψ must be in $PBEST(w,f,g)$. But then $PBEST(w,f,g)$ is not consistent after all: for κ_ϕ entails ϕ and $\neg(\phi \cap \psi)$ and κ_ψ entails ψ and $\neg(\phi \cap \psi)$. This contradicts (1), which entails $PBEST(w,f,g)$ must be consistent.

We also show that epistemic Agglomeration can fail:

Fact 5. (Epistemic Agglomeration Failure) There are w, f, g and Q such that

1. $[\![\text{ought } \phi]\!]^{w,f,g,Q} = 1$ and $[\![\text{ought } \psi]\!]^{w,f,g,Q} = 1$
2. g, the ordering function, is epistemic;
3. and $[\![\text{ought } (\phi \wedge \psi)]\!]^{w,f,g,Q} = 0$

Proof. Simply suppose that ϕ and ψ are consistent but that ϕ and ψ are the only propositions in $PBEST(w,f,g)$. Since g is epistemic, we can stipulate that for no complete answer κ $\kappa \precsim_{w,f,g} \phi$ or $\kappa \precsim_{w,f,g} \psi$.

Finally, consider an important extension of Fact 4. We should also want Agglomeration in the consequent of deontic conditionals:

Conditional Agglomeration. If ϕ, ought ψ, If ϕ, ought $\rho \vDash$ If ϕ, ought $\psi \wedge \rho$

If we adopt a restrictor semantics for conditionals, deontic conditional Agglomeration is a corollary of Fact 3. Say that, where $f + \phi$ is the function such that $f + \phi(w) = f(w) \cap \phi$:

(51) $[\![\text{if } \phi, \psi]\!]^{w,f,g,Q} = 1$ iff $[\![\psi]\!]^{w,f+\phi,g,Q}$=1

With this in hand we can show:

Fact 6. **(No conditional deontic Agglomeration Failure)** There are no w, f, g and Q such that:

1. $[\![\text{If } \rho, \text{ought } \phi]\!]^{w,f,g,Q} = 1$ and $[\![\text{If } \rho, \text{ought } \psi]\!]^{w,f,g,Q} = 1$
2. g, the ordering function, is deontic;
3. and $[\![\text{If } \rho, \text{ought } (\phi \wedge \psi)]\!]^{w,f,g,Q} = 0$

Proof. Given the restrictor semantics, conditional deontic Agglomeration fails just in case there are w, f, g and Q such that:

1. $[\![\text{ought } \phi]\!]^{w,f+\rho,g,Q} = 1$ and $[\![\text{ought } \psi]\!]^{w,f+\rho,g,Q} = 1$
2. g, the ordering function, is deontic;
3. and $[\![\text{ought } (\phi \wedge \psi)]\!]^{w,f+\rho,g,Q} = 0$

But Fact 4 entails these three conditions can never jointly obtain.

References

Luis Alonso-Ovalle. Counterfactuals, correlatives, and disjunction. *Linguistics and Philosophy*, 32(2):207–244, 2009. doi: 10.1007/s10988-009-9059-0.

Fabrizio Cariani. Epistemic and deontic 'should'. *Thought: A Journal of Philosophy*, 2(1):73–84, 2013a.

Fabrizio Cariani. 'Ought' and resolution semantics. *Noûs*, 47(3):534–558, 2013b.

Fabrizio Cariani. Consequence and contrast in deontic semantics. *Journal of Philosophy*, 113(8):396–416, 2016a.

Fabrizio Cariani. Deontic modals and probabilities: One theory to rule them all? In Nate Charlow and Matthew Chrisman, editors, *Deontic Modality*, pages 11–46. Oxford University Press, 2016b.

Jennifer Carr. Subjective *Ought*. *Ergo*, 2, 2015.

Ruth Chang. The possibility of parity. *Ethics*, 112(4):659–688, 2002.

Nate Charlow. Decision theory: Yes! Truth conditions: No! In Nate Charlow and Matthew Chrisman, editors, *Deontic Modals*, pages 47–81. Oxford University Press, 2016.

Stephen Finlay. Oughts and ends. *Philosophical Studies*, 143(3):315–340, 2009.

Kai von Fintel. The best we can (expect) to get? *ms*, 2012.

Kai von Fintel and Sabine Iatridou. How to say 'ought' in foreign: the composition of weak necessity modals. In Jacqueline Guéron and Jacqueline Lecarme, editors, *Time and Modality (Studies in Natural Language and Linguistic Theory 75)*, pages 115–141. Springer, 2008.

Bas Van Fraassen. Values and the heart's command. *Journal of Philosophy*, 70(1): 5–19, 1973.

Melissa Fusco. Deontic modality and the semantics of choice. *Philosophers' Imprint*, 15:1–27, 2015.

Lou Goble. Utilitarian deontic logic. *Philosophical Studies*, 82(3):317–357, 1996.

Simon Goldstein. Free choice and homogeneity. Ms, 2019.

Caspar Hare. Take the sugar. *Analysis*, 70(2):237–247, 2010.

John Hawthorne. Lewis, the lottery and the preface. *Analysis*, 62(3):242–251, 2002.

John F. Horty. *Reasons as Defaults*. Oxford University Press, 2012.

Frank Jackson. On the semantics and logic of obligation. *Mind*, 94(374):177–195, 1985.

Lauri Karttunen. Syntax and semantics of questions. *Linguistics and philosophy*, 1(1):3–44, 1977.

Justin Khoo. Disjunctive antecedent conditionals. *Synthese*, forthcoming.

Angelika Kratzer and Irene Heim. *Semantics in generative grammar*. Blackwells, Oxford, 1998.

Daniel Lassiter. *Measurement and Modality: the Scalar Basis of Modal Semanitcs*. PhD thesis, New York University, 2011.

David Lewis. Ordering semantics and premise semantics for counterfactuals. *Journal of Philosophical Logic*, 10(2):217–234, 1981.

David Lewis. Elusive knowledge. *Australasian Journal of Philosophy*, 74(4):549–567, 1996.

David C. Makinson. The paradox of the preface. *Erkenntnis*, 25(6), 1965.

Matthew Mandelkern. Bounded modality. *Philosophical Review*, 128(1):1–61, 2019.

Matthew Mandelkern, Ginger Schultheis, and David Boylan. Agentive modals. *Philosophical Review*, 126(3):301–343, 2017.

Sarah Moss. On the semantics and pragmatics of epistemic vocabulary. *Semantics and Pragmatics*, 8(1):1–81, 2015.

Dilip Ninan. Two puzzles about deontic necessity. In Jon Gajewski, Valentine Hacquard, Bernhard Nickel, and Seth Yalcin, editors, *New Work on Modality, MIT Working Papers in Linguistics*, volume 51. 2005.

Craige Roberts. Information structure in discourse: Towards an integrated formal theory of pragmatics. *Working Papers in Linguistics-Ohio State University Department of Linguistics*, pages 91–136, 1996.

Alf Ross. Imperatives and logic. *Philosophy of Science*, 11(1):30–46, 1941.

Walter Sinnott-Armstrong. Moral dilemmas and incomparability. *American Philosophical Quarterly*, 22(4):321–329, 1985.

Justin Snedegar. Contrastive semantics for deontic modals. In Martijn Blaauw, editor, *Contrastivism in Philosophy: New Perspectives*. Routledge, London, 2012.

Eric Swanson. The application of constraint semantics to the language of subjective uncertainty. *Journal of Philosophical Logic*, 45(2):121–146, 2016.

Seth Yalcin. Probability operators. *Philosophy Compass*, 5(11):916–937, 2010.

Seth Yalcin. Modalities of normality. In Nate Charlow and Matthew Chrisman, editors, *Deontic Modals*, pages 230–255. Oxford University Press, 2016.