Existential Belief & Embedded Epistemic Modals
Maša Močnik (MIT)

There has been much recent interest in the analysis and distribution of embedded epistemic modals (Yalcin 2007, Anand and Hacquard 2013, a.o.). We present novel data using the embedding verb *dopuščati* ('to allow for the possibility that') from Slovenian, analysed as an existential doxastic attitude, and argue for a new analysis of epistemic modals that captures their restricted distribution under doxastic attitudes.

Data. Anand and Hacquard (2013) (A&H) observe for Romance that universal epistemic modals do not embed under doxastic verbs like *fear/hope/doubt* (analysed as existential), while any force is good under *think/believe* (**B**). Slovenian provides an additional data point with the existential verb *dopuščati* (**D**) (related to Russian *dopuskat'* but different from English *allow*, which is more discursive).

Embedding epistemic modals under *dopuščati*: an existential modal like *utegniti* ($\Diamond$) (similarly for *mogoče* 'perhaps') is somewhat redundant but still acceptable, whereas the universal *morati* ($\Box$) is decidedly odd. Intuitively, this is because it makes *dopuščati* feel inappropriately weak.

(**1**) *Context: John had seen a wet umbrella.*   **B$\Diamond$p**, **D$\Diamond$p**, **B$\Box$p**, **#D$\Box$p**
(a) *Janez {misli/dopušča}, da utegne zunaj deževati.*   (b) *Janez {misli/#dopušča}, da mora$_{ep}$ zunaj deževati.*
  John   thinks allows that might outside to.rain   John   thinks   allows   that must outside to.rain
  John {thinks/allows} it might be raining outside.   John {thinks/#allows} it must$_{ep}$ be raining outside.

Importantly, the embedded readings under inspection are epistemic, with the modal anchored to the attitude holder (not echoing someone else's words or shifted to some other body of evidence). Note that *can't* is also odd under *dopuščati*; this follows nicely from the assumption that *might/can* and *must* are duals ($\Diamond p = \sim\Box\sim p$) and that oddness arises through a wide scope necessity meaning for these items ($\sim\Diamond p$ as $\Box\sim p$).

(**2**) *#Dopuščam, da   ne   more$_{ep}$ biti sončno.*   **#D$\sim\Diamond$p**
    I.allow   that not can   be   sunny
    I allow for the possibility that it can't be sunny.

Unlike A&H's *fear/hope/doubt*, *dopuščati* can be strengthened into a belief claim (reminiscent of *some*, which can be reinforced into *in fact all*), in (3).

(**3**) *Seveda,   dopuščam, da   je Zemlja okrogla – trdno verjamem, da je.*   **Dp&Bp**
    Of course I.allow   that is Earth   round   firmly I.believe that is
    Of course, I allow for the possibility that the Earth is round – I firmly believe that it is.

There is no space to present the context (and anchoring) for negated attitudes, but the judgments pattern differently compared to the positive cases above: (**4**) **$\sim$B$\Diamond$p**, **$\sim$D$\Diamond$p**, **#$\sim$B$\Box$p**, **#$\sim$D$\Box$p**

Embedding epistemic/doxastic attitudes under *dopuščati* is intuitively odd with either force because (unless such an assumption is suspended) agents are assumed to be authorities on their belief/knowledge states.

(**5**) *#Dopuščam, da   {verjamem/dopuščam}, da   zunaj   dežuje.*   **#DDp**, **#DBp**
    I.allow   that I.believe I.allow   that outside it.rains
    I allow for the possibility that {I believe/I allow for the possibility} that it's raining outside.

Dopuščati is thus an embedder with a contrast between $\Diamond$p and Dp (but not $\Box$p and Bp, which are both odd).

Challenge. That modals and attitudes should not get an identical analysis was already pointed out by Yalcin (2007), but it is hard to predict the similarities (#DBp/#D$\Box$p) *and* the differences (#DDp/D$\Diamond$p). (A Yalcin-style revision of epistemic modals (also Mandelkern 2017) is built for oddities arising from the attitude's complement: "embedded epistemic contradictions", and not for those arising from the attitude-modal combination.) Following the footsteps of A&H, there seems to be something more general about existential attitudes that disallows embedded universal epistemic modals. This and the desire to explain #$\sim$B$\Box$p call for an analysis that does not rely on some special feature of *dopuščati* (e.g. a selectional restriction requiring a match in force between D and the modal). A&H's account relies on a special feature of *fear/hope/doubt* that cannot be extended to *dopuščati*: a presupposition on the belief state that requires it to contain some *p* and some non-*p* worlds, where *p* is the prejacent of the embedded modal. This does not fare well with the data in (3), since (3) shows that the doxastic state of *dopuščati* does not need to contain non-*p* worlds. Finally, the relationship between the positive cases in (1) and the negative cases in (4) is complicated: if we "push

the negation in" in (4), we do not obtain parallel judgments to (1) (note: *misliti* 'think' is <u>not</u> neg-raising). Given (2) we want to keep ◊ and □ as duals (contrary to accounts like Ippolito (2017)), so something needs to be said about D and B that sheds light on why D, ~D and ~B do not embed a necessity meaning.

Analysis (of meanings). The first step is to derive the arguably correct meanings of embedded attitudes and epistemic modals. We follow the traditional Hintikka/Kratzer analysis of these as quantifiers over sets of accessible worlds: attitudes specify this set lexically while modals determine their flavour through context. Epistemic modals combine with a free (modal base) variable i (type sst, setting ordering sources aside). We follow Yalcin (2007) in adding an information state parameter κ (a set of worlds) to the index of evaluation (see also Mandelkern (2017) for a generalization of this). <u>Our innovation</u> is to propose that attitude verbs highlight a subset κ' of κ (which we will simply introduce into the index as well): *think/believe* highlight the whole doxastic state, while *dopuščati* highlights the witnesses to its existential statement. Attitudes themselves are not affected by this parameter – they remain duals for non-highlighted content. Symbol A below represents the doxastic agent.

$$[[B_A p]]^{g,κ,κ',w}=1 \text{ iff } \forall w' \in \mathcal{B}_A^W: [\![p]\!]^{g,\mathcal{B}_A^W,\mathcal{B}_A^W,w'}=1 \qquad [[D_A p]]^{g,κ,κ',w}=1 \text{ iff } \exists w' \in \mathcal{B}_A^W: [\![p]\!]^{g,\mathcal{B}_A^W,\{w'\},w'}=1$$

Embedded epistemic modals have received a lot of attention in the philosophical literature in the wake of Yalcin (2007). We follow Mandelkern (2017): epistemic modals have their standard semantics (evidential modal base) but are restricted in not being able to "look outside" of the doxastic state (**Locality**). That is, epistemic modals quantify over subsets of the doxastic state. (This is in place to improve on Yalcin's account, e.g. $K◊p$ no longer follows from p.) Intuitively, one could understand this as assuming that agents do not generally have complete evidence. <u>Our second innovation</u> is to assume that agents actively consider it possible that they have all the evidence: the modal base function maps at least one (highlighted) doxastic world onto the whole doxastic state (**Completeness**).

$[[□_i p]]^{g,κ,κ',w}$ is defined only if $\forall v∈κ: g(i)(v)\subseteqκ$ (Locality) and $\exists v∈κ': g(i)(v)=κ$ (Completeness) and when defined is true iff $\forall w' \in g(i)(w): [\![p]\!]^{g,κ,κ',w'}=1$

$[[◊_i p]]^{g,κ,κ',w}$ is defined only if $\forall v∈κ: g(i)(v)\subseteqκ$ (Locality) and $\exists v∈κ': g(i)(v)=κ$ (Completeness) and when defined is true iff $\exists w' \in g(i)(w): [\![p]\!]^{g,κ,κ',w'}=1$

Example: $[[D_A□_i\, p]]^{g,κ,κ',w}$ is defined only if $\forall v∈\mathcal{B}_A^W: g(i)(v)\subseteq \mathcal{B}_A^W$ and $\exists v∈\{w'\}: g(i)(v)=\mathcal{B}_A^W$ and when defined is true iff $\exists w' \in \mathcal{B}_A^W: \forall w'' \in g(i)(w'): [\![p]\!]^{g,\mathcal{B}_A^W,\{w'\},w''}=1$

Adding Completeness on top of Locality has the effect of strengthening an embedded universal modal but weakening an embedded existential one. The reason for this is that one of the subsets S is extended to the whole doxastic state: if S is required to contain only p-worlds, then the whole doxastic state is now required to contain only p-worlds (whereas the bigger the set S, the weaker the statement that there is a p-world located in S). It thus follows from D□p that the belief state entails p, which is also what follows from B□p. On the other hand, D◊p remains weaker than B◊p. For the negated attitudes, we assume that negation does not interact with the highlighted content (so ~Bφ and D~φ, for example, have different meanings).

Analysis (of oddness). To account for oddness, we use Magri (2009): a sentence is # when the strengthened meaning (via a covert exhaustivity operator *Exh*) gives rise to a contextual contradiction. We assume that D and B are Horn-mates (like *some* and *all*). <u>Doxastics under doxastics</u>: if it is common knowledge that speakers are authorities on their beliefs (e.g. KD45), DDp is # because its pragmatic equivalence to BDp (analogously, DBp$\leftrightarrow_{KD45}$BBp) makes Exh_R(DDp)=DDp&~BDp contradictory. Following Magri (2011) we can restrict the alternatives via R to the minimally relevant ones: exhaust's complement and anything contextually equivalent to it. <u>Embedded epistemic modals</u>: if *Exh* is blind to Locality and Completeness (i.e. if we interpret these as pragmatic principles of how speakers reason about evidence under belief), D□p is # because it becomes equivalent to B□p under Completeness (see above), so Exh_R(D□p)=D□p&~B□p is (pragmatically) contradictory. This matches the intuition that it is odd for the speaker to "only" allow that something *must* be the case. By contrast, Exh_R(D◊p)=D◊p&~B◊p is consistent because D◊p is contextually weaker than B◊p. Similarly, Exh_R(B◊p) and Exh_R(B□p) do not give rise to contextual contradictions (for B◊p we can consider more than what is minimally relevant: Exh_R(B◊p)=B◊p&~B□p). For negated attitudes,

Exh$_R$(~B□p)=~B□p&~~D□p is a contextual contradiction, whereas *Exh$_R$*(~D◊p)=~D◊p and *Exh$_R$*(~B◊p)= ~B◊p&~~D◊p are not. Obligatory embedded exhaustification can be used for #~D□p (see Magri 2011).
Implications. Variation is expected to correlate with how a language (without D) sets up the alternatives.