Weaker-to-Stronger Translational Embeddings in Modal Logic

LLOYD HUMBERSTONE

ABSTRACT. After some general remarks (Section 1), we consider whether a translation popular from earlier discussions of deontic and doxastic logic (reviewed in Section 2) faithfully embeds the normal modal logic **KD** into one of its proper extensions. This would contrast with the more familiar such modal-to-modal embeddings in which a logic is embedded either into one which is strictly weaker or one which is incomparable with it in respect of strength. The translation in question turns out, somewhat surprisingly, not to deliver the desired result (Section 3) and simply in the interests of illustrating the possibility of weaker-to-stronger embeddings amongst the normal modal logics, Section 4 provides an example of a translation which does succeed in embedding **K** (though **KD** could also play this role) into one of its (normal) proper extensions.

Keywords: Translational embeddings, normal modal logics.

1 Introduction

The immediate stimulus for the present study came from some remarks by one of the author's students, Rohan French. We had been discussing the striking paper [2], in which Béziau finds the following situation paradoxical: we have two logics, one stronger than the other (in the sense of proving everything the latter proves, and more besides) while at the same time the stronger logic can be faithfully embedded by a translation into the weaker logic. After all, this embedding locates a copy (under the translation) of the whole of the stronger logic within the weaker – arguably undermining the latter's claim to be weaker. Béziau presents a particularly simple example of this phenomenon in [2], as well as recalling several further instances from the literature. (Still further examples, as well as a discussion of the paradoxicality reaction, may be found in [12].) The following aspect of Béziau's discussion struck French as strange. In *all* the familiar examples of such translational embeddings, a stronger logic is embedded into a weaker one, rather than – what we should expect if Béziau's surprised reaction were appropriate – a weaker logic into a stronger one. (See note 9 for a consideration ignored here.) Indeed, French further reasoned, it would be embedding in this latter direction that would be problematic, since as a general rule a stronger logic collapses more distinctions than a weaker logic, which should be expected to cause trouble for the faithfulness of the supposed translation. A special case of this kind of reasoning was used in the proof of Theorem

3.1 of [12], where it is argued that the consequence relation associated with the $\{\to, \neg\}$-fragment of intuitionistic logic cannot be embedded by (what we call below) a definitional translation into the consequence relation of classical logic on the grounds that for the first of these consequence relations there are, to within equivalence, six non-equivalent formulas in one variable, while for the latter there are only four. (The argument is spelt out in more detail below.) What had not, however, occurred to the author, was that analogous considerations might turn out to threaten *any* translation purporting to embed a weaker logic into a stronger.[1] This was French's suggestion, made in the full knowledge that only a qualified form of it could be expected to survive in view of the following. While it is typically the case that a deductively stronger logic makes fewer discriminations than a weaker one (than a proper sublogic, that is), this is not invariably the case; several types of exception to the general rule are provided in [13].

An explanation is due of some of the terminology in use here. Suppose we think of logics as consequence relations $\vdash_1$ and $\vdash_2$, and regard a *translation* as a function τ from the language of $\vdash_1$ to the language of $\vdash_2$ which is said to *embed* $\vdash_1$ *into* $\vdash_2$ *faithfully* when for all formulas $A_1, \dots, A_n, B$ of the former language, we have:

(1.1) $A_1, \dots, A_n \vdash_1 B$ if and only if $\tau(A_1), \dots, \tau(A_n) \vdash_2 \tau(B)$.

The word 'faithfully' here records the 'if' direction of this condition and since we always impose this restriction, except for occasional emphasis, we simply say that the translation τ *embeds* $\vdash_1$ into $\vdash_2$ when (1.1) is satisfied. These consequence relations are called, respectively, the *source* and the *target* of the embedding in this case. Adapting terminology from previous work on such embeddings,[2] we say in addition that τ is *compositional* when for each k-ary primitive connective # of the language of $\vdash_1$ there is a formula $C(p_1, \dots, p_k)$ of the language of $\vdash_2$ in which precisely the indicated propositional variables (or 'sentence letters') occur – and here we consider only sentential logics[3] – with the property that for all formulas $B_1, \dots, B_k$ of the

[1] [12], p. 162, quotes Ryszard Wójcicki as having said of two logics, one weaker than the other, that *nevertheless* the second can be definitionally embedded in the first, and envisages (p. 163) the following reaction: "There is no 'nevertheless' about it – the increase in expressive power [of the weaker over the stronger logic] is just a manifestation of the general phenomenon that deductively weaker logics exhibit greater discriminatory strength. . . ". [12] then remarks that this response will be seen to fail in a later section of the paper, where (p. 175) what we find is an example of a pair of logics, the stronger of which cannot be embedded in the weaker. The question of current interest, however, is not whether a stronger logic can always be embedded in a weaker one, but rather whether a weaker logic can ever be embedded in a stronger one.

[2] See [11], p. 441, and [12], p. 147, for references. Embeddability – and especially definitional embeddability (introduced presently) – is variously referred to in this literature as interpretability, definability, and reconstructability.

[3] We presume also that the same countably infinite set of such variables $\{p_i\}_{i \in Nat}$ is a set of free generators for the languages of all logics, considered as algebras with their possibly differing sets of primitive connectives as fundamental operations. We generally write p, q, r for p_1, p_2, p_3 respectively. If $A = A(p_1, \dots, p_k)$ is a formula in which the exhibited variables occur, then $A(B_1, \dots, B_k)$ is the result of uniformly substituting B_i

language of $\vdash_1$, $\tau(\#(B_1, \ldots, B_k)) = C(\tau(B_1), \ldots, \tau(B_k))$. τ is *definitional* if it is compositional and we also have $\tau(p_i) = p_i$ for every propositional variable p_i. A certain kind of non-compositional embedding is also familiar from the literature, in which there is a single formula $C(p)$ of the language of $\vdash_2$ such that for any formula B of the language of $\vdash_1$, $\tau(B) = C(B)$. Essentially this amounts to applying a single derived connective from the language of $\vdash_2$ to a formula from the language of $\vdash_2$, and for it to count as a translation we must have the language of $\vdash_1$ included in that of $\vdash_2$. A well-known example is Glivenko's embedding of $\vdash_{CL}$ into $\vdash_{IL}$, where these are the usual consequence relations of classical and intuitionistic propositional logic respectively, and (1.1) holds with $\vdash_{CL} = \vdash_1$, $\vdash_{IL} = \vdash_2$ taking $\tau(B)$ as $\neg\neg B$. (So C is $\neg\neg p$.) In what follows we are concerned mainly with compositional embeddings, and in particular, with the definitional case.

Letting $\vdash_{IL^{\neg}_{\to}}$ and $\vdash_{CL^{\neg}_{\to}}$ be the restrictions of $\vdash_{IL}$ and $\vdash_{CL}$ to the language with connectives $\to, \neg$, let us elaborate upon the point made above (from [12]) that $\vdash_{IL^{\neg}_{\to}}$ cannot be embedded by a definitional translation into $\vdash_{CL^{\neg}_{\to}}$ because the latter makes too many distinctions. More specifically there are six equivalence classes of formulas in the single propositional variable p for $\vdash_{IL^{\neg}_{\to}}$ but only four in the case of $\vdash_{CL^{\neg}_{\to}}$. In general with n distinct such equivalence classes for a consequence relation $\vdash$ there will be $n(n-1)/2$ claims of the form $A_i \dashv\vdash A_j$ with $i < j$, where $A_1, \ldots, A_n$ is an enumeration comprising one representative from each equivalence class. (1.1) implies, for the current choice of $\vdash_1$, $\vdash_2$, and any envisaged τ:

(1.2) $A_i \dashv\vdash_{IL^{\neg}_{\to}} A_j$ if and only if $\tau(A_i) \dashv\vdash_{CL^{\neg}_{\to}} \tau(A_j)$,

where the A_i, A_j, are chosen with $i < j$ (to avoid double-counting) from amongst representatives of the six different $\vdash_{IL^{\neg}_{\to}}$-equivalence classes of formulas in which the only variable to appear is p. There are thus 15 ($= (6\times 5)/2$) such [$A_i \dashv\vdash_{IL^{\neg}_{\to}} A_j$]-claims, all of which are false. Now passing from the left to the right hand side of (1.2), we have the equivalence claims $\tau(A_i) \dashv\vdash_{CL^{\neg}_{\to}} \tau(A_j)$, of which, since there are only four $\vdash_{CL^{\neg}_{\to}}$-equivalence classes to choose from, at most 6 ($= (4 \times 3)/2$) can be true – and perhaps even fewer, since we have not said that τ is surjective (to within equivalence). Thus (1.2) must fail in its 'if' direction for at least 9 ($= 15 - 6$) choices of $\langle i, j\rangle$.[4]

Very little of the hypothesis that τ is a definitional translational embedding from is actually used here in refuting that hypothesis. An examination of the argument establishes, in particular, that there is no τ satisfying the 'if' direction of (1.1) with the property that for all formulas A, $\tau(A)$ is constructed only from variables appearing in A, with $\vdash_1$ as $\vdash_{IL^{\neg}_{\to}}$ and $\vdash_2$ as $\vdash_{CL^{\neg}_{\to}}$. And analogous non-existence results for other weaker-into-stronger embeddings in this vicinity follow by similar reasoning. With full $\vdash_{IL}$ there are even more (namely denumerably many) one-variable formulas, making the shortage of matching distinctions on the $\vdash_{CL}$ side even worse. The same

for p_i ($i = 1, \ldots, k$) in A.

[4] See Theorem 3.1 of [12] and Section 5 of [11] for related observations.

style of argument as has just been given works for the pure ¬-fragments of these consequence relations, as well as for their →-fragments (though in the latter case one must consider instead the two-variable → fragments, since $\vdash_{CL}$ and $\vdash_{IL}$ agree on the one-variable purely implicational formulas, so there is no weaker/stronger situation here).

The above argument about classical and intuitionistic logic ignores the possibility that formulas might be equivalent, each being a consequence of the other, by the consequence relation $\vdash$ at issue, and yet discriminated between ('discernible', 'non-synonymous',...) by $\vdash$ in because of the existence of a contexts in which the replacement of one formula by the other does not yield similarly equivalent results. This is a possibility not arising for $\vdash_{IL}$, $\vdash_{CL}$, and their fragments, but in general it would need to be taken into account in defining the – shall we say – discrimination relation $D_{\vdash}$ of a consequence relation $\vdash$.[5] Suppose we have a set $\mathbb{C}$ of consequence relations on the same language in which the anomalies reviewed in [13] do not arise, and the "strengthening the logic reduces its discriminations" motto applies: i.e., for all $\vdash_0, \vdash_1 \in \mathbb{C}$, we have $\vdash_0 \subsetneq \vdash_1 \ \Rightarrow\ D_{\vdash_1} \subsetneq D_{\vdash_0}$. Does this imply that whenever $\vdash_0 \subsetneq \vdash_1$ (with $\vdash_0, \vdash_1 \in \mathbb{C}$), $\vdash_0$ cannot be definitionally embedded into $\vdash_1$? In fact the earlier argument about the implication-negation fragments of intuitionistic and classical logic does not generalize to give this conclusion. That argument was basically an appeal to the pigeonhole principle: we can't match up the equivalence classes of formulas in one variable in the two logics because there are more such equivalence classes in the intuitionistic case than in the classical case. And what 'more' means in such contexts is that the set of such classes has a greater cardinality: it does not just mean "proper superset".[6] But the latter is all we have from the current hypothesis about stronger logics losing discriminations: $D_{\vdash_1} \subsetneq D_{\vdash_0}$. (A fuller discussion would relativize these D relations further, to the number of variables in the formulas concerned.) If, for example, the two logics both have infinitely many equivalence classes of n-variable formulas, then even if some of these equivalence classes are merged as we pass from the weaker to the stronger logic, a definitional translation which is the identity map except for some n-ary connective, may yet succeed in embedding the weaker logic into the stronger one.

Rather than spend further time on diagnostic remarks, we treat French's idea instead as a challenge: exhibit a concrete example of such an embedding. We pursue this in the arena of modal logic, taking the above n as 1. In Sections 2 and 3, we have a historical look at what come close to being examples, and at logic inspired by them (**KDH**) of considerable interest in its own right. In Section 4 a problem with using that logic as the target of the proposed embedding is overcome, and the translation is changed to

[5] That is, we define: $D_{\vdash}(A, B)$ iff for some formula $C(p)$ of the language of $\vdash$, we have either $C(A) \nvdash C(B)$ or $C(B) \nvdash C(A)$. Here we do *not* require that $C(p)$ have p for its only propositional variable. (Here we are considering how to make allowance for non-congruential – also called "non-selfextensional" – consequence relations.)

[6] Interestingly, 'Galileo's Paradox' is also mentioned in Béziau [2], though not quite in the present connection.

something of less independent interest but of great service in meeting the challenge just described.

2 Turning to Modal Logic

We will consider modal logics for simplicity as sets of formulas rather than as consequence relations, and as all cast in the same language with any functionally complete set of primitive boolean connectives (taken as fixed throughout, and preferably containing the nullary truth connective $\top$) together with the 1-ary connective $\Box$ ($\Diamond A$ being defined as $\neg\Box\neg A$). A modal logic $\mathbf{S}$ is any set of formulas in this language which contains all truth-functional tautologies and is closed under uniform substitution (of arbitrary formulas for propositional variables) and Modus Ponens. A (not necessarily primitive) 1-ary connective $\#$ is *normal* in $\mathbf{S}$ if $A \to B \in \mathbf{S}$ implies $\#A \to \#B \in \mathbf{S}$ (all A, B), $(\#p \wedge \#q) \to \#(p \wedge q) \in \mathbf{S}$, and $\#\top \in \mathbf{S}$. Finally, a modal logic $\mathbf{S}$ is *normal* if $\Box$ is normal in $\mathbf{S}$. The reasoning for defining the notion of a normal modal logic in this way will be evident as we proceed. Its upshot coincides with that of the usual definition – though for some purposes it is not convenient to build in the condition of closure under uniform substitution that has been imposed here. In future rather than writing "$A \in \mathbf{S}$" we write "$\vdash_{\mathbf{S}} A$", and describe A as provable in $\mathbf{S}$. (The above definition of normality evidently gives rise to an axiomatization of the smallest normal modal logic $\mathbf{K}$, and likewise with other logics isolated, so provability here amounts to the existence of a proof from the axiomatization in question. Familiarity with reasonably standard names of various modal logics and modal formulas is assumed in what follows.) Given this use of "$\vdash$", the definitions of embedding etc., from the preceding section can be transferred to the present setting but with $n = 0$ in (1.1).[7] Several non-compositional embeddings in the ('one-off prefixing') style of Glivenko's, mentioned in that section, are known: for example –

- $\vdash_{\mathbf{S5}} A$ if and only if $\vdash_{\mathbf{S4}} \Diamond\Box A$ (Matsumoto [19]);
- $\vdash_{\mathbf{S5}} A$ if and only if $\vdash_{\mathbf{K5}} \Box A$ (Chellas [6]).

Of interest to us here, however, are compositional – indeed definitional – embeddings, and since we are discussing modal logics, the languages of the source and target logics for these embeddings coincide. (The contrast here is with modal translations of intuitionistic, intermediate and substructural logics. See [4], [9]; these are typically compositional though not definitional translations.) Matters are further simplified if we restrict attention to those translations τ which translate the boolean connectives by themselves, in

[7] We could of course pursue the modal discussion in terms of consequence relations also, for which purpose the most useful consequence relation to associate with $\mathbf{S}$ would be the 'local' consequence relation holding between a set Γ of formulas and a formula A when a conjunction of formulas from Γ provably implies A in $\mathbf{S}$. There seems little advantage in doing this, though. The discernibility relations $D_{\mathbf{S}}$ relate A to B if for some formula $C(p)$, $\vdash_{\mathbf{S}} C(A)$ while $\nvdash_{\mathbf{S}} C(B)$ or vice versa, which in turn holds for normal modal logics $\mathbf{S}$ just in case $\nvdash_{\mathbf{S}} A \leftrightarrow B$.

the sense that for all primitive connectives # (of arity k) other than $\Box$, we have $\tau(\#(B_1,\ldots,B_k)) = \#(\tau(B_1),\ldots,\tau(B_k))$. A familiar example of a definitional embedding along these lines is given by the translation τ for which $\tau(\Box B) = \tau(B) \wedge \Box\tau(B)$, which embeds **KT** into **K**. This example is discussed in [20], where the question is raised of a similar translation in the reverse direction, embedding **K** into **KT**. This last question was answered negatively in [12], p. 174, with the observation that no translation which translates a formula into one in the same propositional variables and translates $\vee$ (disjunction) by itself can embed a Halldén-incomplete logic into a logic which is Halldén-complete.[8] This may seem a somewhat parochial feature of the relationship between **K** and **KT** in view of the considerations of Section 1: the more glaring aspect of this relationship is that $\mathbf{K} \subsetneq \mathbf{KT}$ so the envisaged embedding would be in the contentious "weaker into stronger" direction. Note incidentally that each of the two non-compositional examples listed above again embeds a stronger logic into a weaker one. Evidently we shall need to look elsewhere if we are to find a weaker-to-stronger embedding between modal logics.[9]

Where else to look, then? We can usefully cast our minds back to 1959 for an early pointer, with the publication of Dawson [8], in which the problem of representing a suitable deontic logic within the expressive resources of a suitable alethic modal logic.[10] Here the main features of suitability are that in the alethic case, we should have the formula **T** ($= \Box p \to p$) provable, whereas in the deontic case we should not, asking only for the provability of the weaker **D** (i.e., $\Box p \to \Diamond p$, or equivalently, if we are working amongst the normal modal logics, $\Diamond\top$), and perhaps also the formula called **U** in [17]: $\Box(\Box p \to p)$. Thus our deontic $\Box$ operator (often written "O" in such discussions) can't just be the alethic $\Box$ operator; but Dawson noticed that if we took the deontic $\Box$ as, instead, the alethic $\Diamond\Box$, and took **S4.2** as our alethic modal logic, the normal extension of **S4** by the formula **G**, as it sometimes known (after Geach), then a more suitable deontic logic emerges. We retain the conventional label **S4.2** rather than the explicit name **S4G** (or indeed, more explicitly still, **KT4G**) for this logic, in the interests of familiarity:

G $\quad \Diamond\Box p \to \Box\Diamond p$

then the prospects are rosier, since while the unwanted "**T**-for-$\Diamond\Box$" formula remains unprovable, the prefix $\Diamond\Box$ is normal in **S4.2** (though not in **S4**) – *cf.* note 15 – and **G** itself is the required deontic weakening "**D**-for-$\Diamond\Box$"

[8] For another application of this fact, see note 19 of [15].

[9] Nothing in this formulation should make us ignore the third kind of case: embedding one logic into another which is neither stronger nor weaker than it. It is easy to find modal examples illustrating this possibility. For instance take the logic determined by the class of all frames when truth at a point in a model is defined in such a way that $\Box A$ is taken to be true at a point just in case A is not true at any accessible point. This modal logic – not a normal modal logic, to be sure – is $\subseteq$-incomparable with **K**, but can be embedded in **K** by the translation of $\Box A$ as $\Box\neg A$. We will see an example of the same phenomenon amongst the normal modal logics in the last paragraph of this section.

[10] Variations on Dawson's theme may be found in Åqvist [1] and Kielkopf [16].

("*Ought* implies *May*") of that unwanted principle. (Note that when '$\Box$' is translated into '$\Diamond\Box$', '$\Diamond$' becomes '$\Box\Diamond$'.)[11]

Twenty years later, this same alethic/deontic issue came up as an epistemic/doxastic issue, with the problem of representing a suitable doxastic $\Box$ in terms of an epistemic $\Box$. Again we may want **D** for the former (belief consistency as part of our idealization), and perhaps **U** as well – *cf.* note 42 in [14] – but we certainly don't want **T**, and Wolfgang Lenzen, in [18],[12] again reached for **S4.2** as a suitable epistemic logic, and the suggestion that the belief operator (our doxastic $\Box$) be defined in terms of the epistemic $\Box$ (often written as K in this context), namely as $\Diamond\Box$ again. Lenzen

[11]**T** here is more than we need: $\Diamond\Box$ is easily seen to be normal in any extension of **KD4.2**. In particular, its normality in **KD4.3** was noted – though not described in these terms – in Rescher and Urquhart [22], pp. 135–137, where it is, in effect, observed that the embedding which replaces $\Box$ with $\Diamond\Box$ embeds **KD5** into **KD4.3**, in the sense of 'embed' not requiring faithfulness, and it is conjectured that the embedding is indeed faithful. Byrd [3], p. 592, tries to establish this conjecture by appeal to a result from Segerberg [24] to the effect that the only consistent proper extensions of **KD5** (called *D5* in [3]) are **S5**, **KD5Alt**$_n$, and **S5Alt**$_n$ ($n \in Nat$); but there is no such result to be found in the pages cited by Byrd from [24], and it would not be correct since most conspicuously missing from this list is **KD45** (together with its **Alt**$_n$ extensions limiting the numbers of points accessible to a given point). The unprovability of **4** in **KD5** shows in fact that Rescher and Urquhart's conjecture is false, since its translation is a theorem of **KD4.3** (indeed of **KD4**). (Whether, revised so as to have **KD45** as a source logic for the translation, the conjecture would be correct, the author does not know.) It should be pointed out that in the works cited, none of this terminology is employed. Instead, Rescher and Urquhart observe that for the (future-directed) tense logic of linear unending time if we interpret $\Box$ as what in Prior's notation would be written as FG, the logic **KD5** (for $\Box$) is sound, and conjecture that it is complete. (Soundness and completeness here are to be understood as "with respect to the class of (forward) serial, transitive, and (forward) connected frames".) Byrd took himself to be establishing that this logic is indeed complete for the intended interpretation of $\Box$ ("eventual permanence") over these frames, and point just made would be put by saying that $\Box p \to \Box\Box p$ is valid but not provable, contradicting the supposed completeness result.

[12]Not being a reader of German, I owe my knowledge of the contents of this paper to correspondence from Lenzen in 1982; the parallel with Dawson occurred to me only somewhat later. A significant difference between the two – which does not bear on the issue of embeddings – is that Lenzen is also concerned with a plausible *mixed* epistemic-doxastic logic in which, to use the conventional notations for the $\Box$ operators in this setting, BA is defined to be $\neg K \neg KA$. Thus there are questions about the interaction of knowledge and belief to consider. The most interesting interaction effect is a 'Cartesian' principle: $Bp \to BKp$, reflecting the fact that belief here is taken to be of the 'fully confident' variety. Dawson, on the other hand, was not concerned in [8] with developing a plausible combined alethic-deontic logic – except on the second last page of [8], where he ventures into this territory as a kind of afterthought. There is, for instance, no need to defend the analogue of the mixed epistemic-doxastic principle just cited: that whatever ought to be the case ought necessarily to be the case. Another way of putting this point is to say that whereas Lenzen is interested in the definitional extension of an epistemic logic by the definition of B given above, Dawson (or an idealized version of Dawson, ignoring the afterthought) is interested only in the image under the correlated τ – in the present instance, that replacing $\Box$ with $\Diamond\Box$ – of a given alethic logic. (Note that it is because of the correlation between these embeddings and explicit definitions that the former are called definitional embeddings; when used as explicit definitions, however, it is mandatory that the $\Box$ in the *definiens* and the $\Box$ in the *definiendum* be notationally distinguished, as with the K-versus-B notation earlier.)

further observed that the (definitional) translation which simply replaces $\Box$ with $\Diamond\Box$ throughout a formula embeds **KD45** (a suitably idealized doxastic logic – sometimes also referred to as 'deontic **S5**'[13] into **S4.2** (a candidate epistemic logic).[14]

Although the above description represents the construction of a deontic or doxastic logic on the basis of an alethic or epistemic logic as a process of weakening (weakening **T** to **D**), when we look more closely, we can see that there is also some strengthening going on. This is evident from the fact that the source of the embedding just mentioned is **KD45**, which because of the '**5**' is not after all a sublogic of its target, **S4.2**. So we are not quite in possession of an example of the phenomenon of interest: a weaker-to-stronger definitional embedding. (This example may be added to that in note 9 to illustrate the case of incomparable source and target logics, this time from amongst the normal modal logics.) Still fixing on the translation that replaces $\Box$ by $\Diamond\Box$ we could try correcting matters by either strengthening the target logic – say, to **S5**, since the source logic proves **5** – or by weakening the source logic, to something in which **5** in particular, is not provable. The former avenue is in fact not open because the fact that $\nvdash_{\mathbf{KD45}} \Box p \to p$ would require $\nvdash_{\mathbf{S5}} \Diamond\Box p \to p$. So in the following section we pursue the latter possibility.

3 The Logic KDH

The most important feature of **S4.2** in explaining its appearance as the target logic for the embedding mentioned in the previous section is not the leading axiom **G** itself, but the fact that in concert with the rest of **S4**, this logic includes amongst its theorems the formula we shall call **H**, simplifying some given in in Chellas [7] to two of the formulas listed here;[15] we use the label indiscriminately for any of the following four following evidently interdeducible forms:

[13] This same label has been used for several different logics. Even as early as 1971, Rescher and Urquhart [22], p. 136 (n. 14 and text to which the note is appended), cite three such logics.

[14] And further still, Lenzen notes that this remains correct if the target is replaced by the stronger logic **S4.4**, for which case there is also a translation in the reverse direction. This aspect of Lenzen [18] is recalled in Pelletier and Urquhart [21]. Interestingly, in the final paragraph of [8], Dawson had also mentioned **S4.4**, though without using that name.

[15] The labels in question are "($\mathbf{H}^{++}$)" and"($\mathbf{H}^{++}\Diamond$)". We have taken the liberty of removing the "++" superscript, even though another formula on Chellas's list, in Exercise 4.56 ([7], p. 144) has this label (i.e., **H** *tout court*), and – contrary to the claim there made – the result of adding the present formula to **S4** is not **S4.3** but **S4.2**. The result of adding the formula which *Chellas* calls **H** (following, e.g., [17], p. 69, and named after Hintikka) is indeed **S4.3**, and the list of formulas in the exercise falls into two parts, some of which give **S4.2** and others of which give **S4.3**, when added to **S4**. The needed correction seems not to be amongst the several which have been made in subsequent reprintings of [7] – at least, this is so for the 1995 printing. Yet another (unrelated) formula is called **H** in Hughes and Cresswell [10], p. 160; likewise in the cases of Segerberg [24], p. 148, and Chagrov and Zakharyaschev [5], Exercise 5.21, p. 158.

$$\mathbf{H} \qquad (\Diamond\Box p \wedge \Diamond\Box q) \to \Diamond\Box(p \wedge q) \qquad (\Diamond\Box p \wedge \Diamond\Box q) \to \Diamond(\Box p \wedge \Box q)$$
$$\Box\Diamond(p \vee q) \to (\Box\Diamond p \vee \Box\Diamond q) \qquad \Box(\Diamond p \vee \Diamond q) \to (\Box\Diamond p \vee \Box\Diamond q).$$

While **H** and **G** yield the same normal modal logic (namely **S4.2**) as additions to **S4**, for purposes of extending **K** (and **KD**) it is the effects of **H** that we need. (**G** is easily seen to be provable in **KDH**.)[16]

It will also be convenient to isolate the formulas $\mathbf{H}_n$, though this time we mention only the analogues of the first formula listed under **H**:

$$\mathbf{H}_n \qquad (\Diamond\Box p_1 \wedge \ldots \wedge \Diamond\Box p_n) \to \Diamond\Box(p_1 \wedge \ldots \wedge p_n).$$

it is easy enough to see that any normal modal logic proving $\mathbf{H}_2$ ($= \mathbf{H}$) proves $\mathbf{H}_n$ for all $n \geq 1$. Since the converse is even more easily seen, we have $\mathbf{KH}_m = \mathbf{KH}_n$ for any $m, n \geq 2$. And $\mathbf{KH}_1 = \mathbf{K}$.[17] It also easy to check that **H** modally defines the class of frames $\langle W, R\rangle$ satisfying a condition[18] we shall call H, taking $\langle W, R\rangle$ as a first-order structure:

$$\mathsf{H} \qquad \forall w, x, y((Rwx \wedge Rwy) \to \exists u(Rwu \wedge \forall v(Ruv \to (Rxv \wedge Ryv)))$$

and a familiar canonical model argument shows that **KH** is determined by (= sound and complete w.r.t.) the class of frames satisfying H. The key step, recapitulated in the Appendix for the sake of a contrast with a variant mentioned immediately below, involves showing that the canonical frame for any (consistent) normal modal logic extending **KH** satisfies H. For present purposes what is of greater interest is that the canonical frame for any such logic satisfies a stronger condition, H^+, from which it follows that **KH** is determined not only by the class of frames satisfying H but also by the narrower class of frames satisfying H^+. (See the Appendix for a proof of this.)

$$\mathsf{H}^+ \qquad \forall w[\exists u(Rwu) \to \exists u(Rwu \wedge \forall x(Rwx \to \forall v(Ruv \to Rxv)))]$$

Why are we taking all this interest in the axiom **H**? Not because of the logic **KH**, but because of the logic **KDH**, in fact. Recall that for a primitive or derived 1-ary connective # to be normal in **S** we require for all formulas A, B: that (i) whenever $\vdash_{\mathbf{S}} A \to B$ we have $\vdash_{\mathbf{S}} \#A \to \#B$, (ii) $\vdash_{\mathbf{S}} (\#A \wedge \#B) \to \#(A \wedge B)$, and (iii) $\vdash_{\mathbf{S}} \#\top$. Now in any normal modal logic condition (i) is satisfied by any affirmative modality, and hence by the choice of # as $\Diamond\Box$, and condition (ii) is equivalent, given the requirement of closure under uniform substitution, to the provability of **H**; but condition (iii) for the present choice of # is not satisfied by **KH**, since $\nvdash_{\mathbf{KH}} \Diamond\Box\top$. Thus we need to consider an extension of **KH**, and since in any normal modal logic $\Diamond\Box\top$ and $\Diamond\top$ are equivalent (because $\Box\top$ and $\top$ are), the extension we need is **KDH** – alias the smallest normal modal logic containing $\mathbf{H}_n$ for all all $n \geq 0$ (as we saw in note 17).

[16] The interdeducibility of **G** and **H** over **S4** is shown on p. 75 of [8]; this is part of the point of Chellas's exercise, mentioned in the previous footnote.

[17] What about $\mathbf{H}_0$? Understanding the conjunction with no conjuncts as $\top$, the antecedent can be ignored, leaving us with $\Diamond\Box\top$, which is, as noted below, equivalent to **D**.

[18] That is, **H** is valid on all and only frames satisfying this condition.

The considerations just rehearsed show not just that $\Diamond\Box$ is normal in **KDH**, but that **KDH** is the *smallest* normal modal logic in which $\Diamond\Box$ is normal. This means that all the behaviour demanded of $\Box$ for the logic to count as normal is displayed by $\Diamond\Box$ in **KDH** but in no weaker logic. But what has still not been ruled out is that $\Diamond\Box$ might, considered as a single operator, exhibit other logical behaviour beyond that demanded by normality (i.e., exhibited by $\Box$ in **K**). Ruling that out would constitute establishing, that the Dawson–Lenzen translation – call it $\tau_{\Diamond\Box}$ – which uniformly replaces $\Box$ with $\Diamond\Box$ does indeed embed **K** into **KDH**. We can see without much difficulty that this can't be quite right, since any extension of **KD** proves $\Box\Diamond\top$ and this is $\tau_{\Diamond\Box}(\Diamond\top)$ while $\Diamond\top$ is not **K**-provable. Thus, we need to replace **K** as a candidate source logic by **KD**, and adjust the conjecture to: $\tau_{\Diamond\Box}$ embeds **KD** into **KDH**. Spelling this out explicitly:

KDH Conjecture: For any formula A: $\vdash_{\mathbf{KD}} A \Leftrightarrow\ \vdash_{\mathbf{KDH}} \tau_{\Diamond\Box}(A)$.

The $\Rightarrow$ direction follows from the normality of $\Diamond\Box$ in **KDH** and the **KDH**-provability of $\tau_{\Diamond\Box}(\mathbf{D})$. This amounts to a proof of the $\Rightarrow$-claim by induction on the length of (a shortest) proof of A in **KD**. Before addressing the converse direction, we pause to ponder one aspect of the current direction.

The normality of $\Diamond\Box$ in **KDH** suggests that it should be interpretable by universal quantification over accessible points for some suitable accessibility relation. Making this explicit gives an alternative proof of the $\Rightarrow$ direction of the **KDH** conjecture, but seems to be of some interest in its own right, as one would like to see this implicit binary relation exhibited. As a preliminary step, we need a general observation concerning frames satisfying a strengthening of the condition H^+ above, which we call DH^+.

DH^+ $\qquad \forall w\exists u(Rwu \wedge \forall x(Rwx \to \forall v(Ruv \to Rxv)))$

Note that DH^+ is equivalent to the conjunction of H^+, given earlier, with the 'seriality' condition $\mathsf{D} = \forall w\exists z(Rwz)$.

For any w, call any point u as promised by DH^+ a *special successor* of w. Thus a special successor of w is an R-successor of w which, for any R-successor x of w, bears the relation R only to such points as x bears the relation R to. (The original condition H^+ says only that any point with a successor has a special successor.) To fix notation: we regard a model as a structure $\langle W, R, V\rangle$, where $\langle W, R\rangle$ is the frame of the model – $W \neq \varnothing$ and $R \subseteq W \times W$ – and V is a function mapping propositional variables to subsets of W. Truth at a point in such a model is then taken as defined in the standard way for arbitrary formulas, and notated with the aid of "$\models$".

PROPOSITION 1. *Let $\mathcal{M} = \langle W, R, V\rangle$ be any model whose frame $\langle W, R\rangle$ satisfies DH^+. Thus every $w \in W$ has a special successor in the sense just defined. Arbitrarily choose one such special successor w^* for each $w \in W$ and define a new model $\mathcal{M}^* = \langle W, S, V\rangle$ where Sxy just in case Rx^*y. Then for every formula A: for all $w \in W$ we have $\mathcal{M}^* \models_w A$ if and only if $\mathcal{M} \models_w \tau_{\Diamond\Box}(A)$.*

Proof. By induction on the complexity of A. The only case of interest arises in the induction step, and concerns the case in which A is $\Box B$ for some formula B. Since $\tau_{\Diamond\Box}(\Box B)$ is $\Diamond\Box(\tau_{\Diamond\Box}(B))$, what we have to show is that:

$$\mathcal{M}^* \models_w \Box B \text{ if and only if } \mathcal{M} \models_w \Diamond\Box(\tau_{\Diamond\Box}(B)).$$

For the 'only if' direction, suppose that $\mathcal{M}^* \models_w \Box B$. Then for all $v \in W$ such that Swv, we have $\mathcal{M}^* \models_v B$. This means that for all v such that Rw^*v, $\mathcal{M}^* \models_v B$. Suppose, for a contradiction that $\mathcal{M} \not\models_w \Diamond\Box(\tau_{\Diamond\Box}(B))$. In particular then, $\mathcal{M} \not\models_{w^*} \Box(\tau_{\Diamond\Box}(B))$, so for some $x \in W$ for which Rw^*x, $\mathcal{M} \not\models_x \tau_{\Diamond\Box}(B)$ and thus by the inductive hypothesis, $\mathcal{M}^* \not\models_x B$. Now we have our contradiction with the earlier conclusion that for all v such that Rw^*v, $\mathcal{M}^* \models_v B$.

For the 'if' direction, suppose that (i) $\mathcal{M} \models_w \Diamond\Box(\tau_{\Diamond\Box}(B))$ but (ii) $\mathcal{M}^* \not\models_w \Box B$. From (ii) we infer that for some $v \in W$, Swv, which is to say: Rw^*v, and $\mathcal{M}^* \not\models_v B$. From (i), there is $x \in W$ with Rwx and $\mathcal{M} \models_x \Box(\tau_{\Diamond\Box}(B))$. Since w^* is a special successor of w and Rw^*v and Rwx, we must have Rxv (as $\mathcal{M}$ satisfies DH^+); so, since $\mathcal{M} \models_x \Box(\tau_{\Diamond\Box}(B))$, $\mathcal{M} \models_v \tau_{\Diamond\Box}(B)$. The inductive hypothesis then gives: $\mathcal{M}^* \models_v B$, contradicting what was inferred from (ii) above. ■

The S appearing here is the 'hidden' accessibility relation which provides a model-theoretic explanation of the normality of $\Diamond\Box$ in **KDH**. To use this observation to derive the (already settled) "$\Rightarrow$" direction of the **KDH** Conjecture, suppose that $\nvdash_{\mathbf{KDH}} \tau_{\Diamond\Box}(A)$. Thus A is false as some point w the canonical model for **KDH**, which satisfies the condition DH^+ (see Appendix). Thus the construction used in Proposition 1 gives another model (with the same universe) in which A itself is false at w. Further, this new model has a serial frame. Therefore, by the soundness of **KD** w.r.t. the class of such frames, $\nvdash_{\mathbf{KD}} A$.

Let us return to the $\Leftarrow$ direction of the **KDH** Conjecture above. By analogy with the case of Proposition 1, one would like a recipe for turning any serial model into a model for **KDH**, perhaps adding additional points (to serve as special successors) in such a way that whenever a point in the first model falsifies a formula A, a corresponding point in the second model falsifies $\tau_{\Diamond\Box}(A)$. That would guarantee that whenever **KDH** proves the $\tau_{\Diamond\Box}$-translation of a formula, the formula itself is **KD**-provable. But, perhaps surprisingly in view of the fact, emphasized above, that **KDH** is the *weakest* normal modal logic in which $\Diamond\Box$ is normal, it is too strong for any such recipe to be found:

PROPOSITION 2. *The $\Leftarrow$ direction of the* **KDH** *Conjecture is false.*

Proof. It suffices to provide a counterexample – a formula A for which $\vdash_{\mathbf{KDH}} \tau_{\Diamond\Box}(A)$ while $\nvdash_{\mathbf{KD}} A$. We do so by selecting $A = (\Diamond\Diamond p \wedge \Diamond\Diamond q) \to \Diamond(\Diamond p \wedge \Diamond q)$. There is no difficulty in falsifying this formula at a point in a model on a serial frame, so it is not **KD**-provable. But its $\tau_{\Diamond\Box}$-translation is

KDH-provable, as the following considerations show. Begin with the upper right-hand form of the **H** axiom above:

$$(\Diamond\Box p \wedge \Diamond\Box q) \to \Diamond(\Box p \wedge \Box q).$$

By the normality of $\Box$, we conclude that the following is provable:

$$(\Box\Diamond\Box p \wedge \Box\Diamond\Box q) \to \Box\Diamond(\Box p \wedge \Box q).$$

Finally, substituting $\Diamond p$ for p and $\Diamond q$ for q:

$$(\Box\Diamond\Box\Diamond p \wedge \Box\Diamond\Box\Diamond q) \to \Box\Diamond(\Box\Diamond p \wedge \Box\Diamond q).$$

We have here arrived at $\tau_{\Diamond\Box}(A)$ for the choice of A indicated above. ■

Earlier, we noted that the source logic for the conjectured embedding had to be boosted from **K** to **KD** because $\tau_{\Diamond\Box}(\mathbf{D})$ was **KDH**-provable, which was no problem for the general prospect of a weaker-to-stronger embedding, since **KD** $\subsetneq$ **KDH**. The above example does not allow for a similar 'minor adjustment', as the formula $(\Diamond\Diamond p \wedge \Diamond\Diamond q) \to \Diamond(\Diamond p \wedge \Diamond q)$ is not provable in **KDH**. Because of this negative finding, we conclude our exploration of the translation $\tau_{\Diamond\Box}$ at this point, and despite the interest, in view of the observations of Dawson and Lenzen, of this translation turn our attention to finding a weaker-to-stronger embedding (via a definitional translation) from whatever quarter we may, setting aside such questions as whether the translation concerned is of independent interest.

4 A Successful Weaker-to-Stronger Embedding

We will make use of the idea, mentioned before Proposition 1, of a model construction adding a new layer of points between points and those accessible to them in some given model, so that what took one step by the accessibility relation now takes two steps by the new relation. This was problematic in the case where we wanted the new model to be a model for **KDH**, because such conditions as H and H^+ had to be satisfied by the new model – which Proposition 1 shows is a requirement that could not be satisfied. We will now consider the translation $\tau_{\Box\Box}$ which replaces every $\Box$ by $\Box\Box$. The latter is normal in every normal modal logic, allowing us to bypass the difficulties (just alluded to) raised by **H**, whose presence was required in the target logic for $\Diamond\Box$ to be normal. We start with a model $\mathcal{M} = \langle W, R, V\rangle$ and construct what we shall call $\mathcal{M}^+$. (To visualize this process, it is simplest to imagine that the frame of $\mathcal{M}$ is a tree, restriction to which cases will be sufficient for our applications of the idea, in view of [23].) For each element of W pick a new element $\bar{w}$ where $\bar{w} \notin W$ and for $w, v \in W$, $\bar{w} \neq \bar{v}$ whenever $w \neq v$. Now define $\mathcal{M}^+$ to be $\langle W^+, S, V^+\rangle$ where:

- $W^+ = W \cup \{\bar{w} \,|\, w \in W\}$;

- for $x, y \in W^+$, Sxy if and only if for some $w \in W$ we have $x = w$ and $y = \bar{w}$ or else $x = \bar{w}$, $y \in W$ and Rwy;
- V^+ is any assignment of truth-sets to atomic formulas whose restriction to W is V. (I.e., for all $w \in W$, $w \in V^+(p_i) \Leftrightarrow w \in V(p_i)$; this means we have not uniquely specified $\mathcal{M}^+$, but may take it to be any model meeting the constraints given.)

LEMMA 3. *Let $\mathcal{M} = \langle W, R, V\rangle$ be any model and take $\mathcal{M}^+$ to be the model $\langle W^+, S, V^+\rangle$ obtained by the construction just described. Then for all formulas A, and all $w \in W$ we have $\mathcal{M} \models_w A$ iff $\mathcal{M}^+ \models_w \tau_{\Box\Box}(A)$.*

Proof. As for Proposition 1, this is shown by induction on the complexity of A and the only inductive case of interest is again that of $A = \Box B$, which is much simpler here as Rwv (for $w, v \in W$) iff $Sw\bar{w}$ and $S\bar{w}v$, which in turn holds iff S^2wv. ■

Lemma 3 allows us to draw the not very exciting conclusion that $\tau_{\Box\Box}$ embeds **K** into itself. The '$\Rightarrow$' half of this conclusion:

$$\text{For all } A, \vdash_{\mathbf{K}} A \;\Leftrightarrow\; \vdash_{\mathbf{K}} \tau_{\Box\Box}(A),$$

is something already mentioned in passing, that $\Box\Box$ is normal in any normal modal logic, and so in particular in **K**, which itself, like the corresponding claim for $\Diamond\Box$ in our discussion of **KDH** in the preceding section, can be shown either by an induction on the length of proofs in an axiomatization (in the present instance, of **K**), or model-theoretically, as in the proof of Proposition 1: in the present case this amounts to the observation that the relative product R^2 of a relation R with itself, is just another binary relation – which can be used to interpret $\Box$. The '$\Leftarrow$' half of the inset claim, presented contrapositively, invites us to consider a point in a model $\mathcal{M}$ at which A is false (by the completeness of **K** w.r.t. the class of all frames), from which Lemma 3 allows us to infer that $\tau_{\Box\Box}(A)$ is false at the corresponding point in the model $\mathcal{M}^+$, and is therefore not **K**-provable, this time by the soundness of **K** w.r.t. the class of all frames.

To get some mileage out of Lemma 3 for where we have been heading, however, we need to observe that the frames of the derived models $\mathcal{M}^+$ have some specific structural features that can be exploited to validate modal formulas not valid on all frames. Specifically, note that each point in the universe W^+ of such a frame is either one of the original elements w of W, in which case it bears the new accessibility relation S to at most one point, namely the 'new' point $\bar{w}$, or else it is itself one of these new ("barred") points, in which case all its S-successors in turns bear the relation S to at most one point. Thus the following formula, for which we adopt an *ad hoc* notation simply in the interest of brevity:

X $\quad (\Diamond p \to \Box p) \vee \Box(\Diamond q \to \Box q)$

is valid at each point in the frame, as the first first disjunct is valid at the surviving 'old' points and the second at the new points.[19] Where **KX** is the smallest normal modal logic containing the above formula **X**, we can combine this 'soundness' observation with a completeness result provided straightforwardly by the canonical model method to record the following, stated with a reversion to the 'R' notation for an arbitrary accessibility relation and also write $R(w)$ for $\{z \in W \,|\, Rwz\}$, etc. (and $|Z|$ is the cardinality of a set Z):

PROPOSITION 4. **KX** *is determined by the class of frames* $\langle W, R\rangle$ *in which for all* $x \in W$ *either* $|(R(x))| \leq 1$ *or for all* $y \in R(x), |R(y)| \leq 1$.

All we need for present purposes is the soundness half of this result, though, already evident in the observation that **X** is valid on the frames of the models $\mathcal{M}^+$.

THEOREM 5. *The translation* $\tau_{\Box\Box}$ *embeds* **K** *(faithfully) into* **KX**.

Proof. We have already seen that $\tau_{\Box\Box}(A)$ is **K**-provable – so *a fortiori* it is **KX**-provable – for any **K**-provable A. On the other hand if $\nvdash_{\mathbf{K}} A$, then for some model $\mathcal{M}$ and some point w in that model, $\mathcal{M} \not\models_w A$ so by Lemma 3, $\mathcal{M}^+ \not\models_w \tau_{\Box\Box}(A)$, and since $\mathcal{M}^+$ is a model for **KX**, $\nvdash_{\mathbf{KX}} \tau_{\Box\Box}(A)$. ■

Note incidentally, *à propos* of the observation in Section 1 about the unembeddability of Halldén-incomplete logics, such as **K**, into Halldén-complete logics, that **KX** is evidently Halldén-incomplete, in view of the 'Halldén-unreasonable' disjunction **X**. (Use Proposition 4 to check that neither disjunct is **KX**-provable.)

The same considerations show that $\tau_{\Box\Box}$ embeds **KD** into **KDX**, because the transformation $\mathcal{M} \rightsquigarrow \mathcal{M}^+$ preserves seriality. (For the analogue of Proposition 4 in the case of **KDX**, replace the occurrences of "$\leq$ with "$=$"; the "$\rightarrow$"s in **X** can also be equivalently replaced with "$\leftrightarrow$"s in the presence of **D**.) It would be interesting to know what range is of normal modal logics for which an argument along these lines can be given.

It is also worth recording the fact that the transformation from $\mathcal{M}$ to $\mathcal{M}^+$ also allows us to prove the analogue of Lemma 3 with $\tau_{\Box\Box}$ replaced by the translation that occupied us in Section 3, $\tau_{\Diamond\Box}$, and that for all $w \in W$, $\mathcal{M}^+ \models_w A$ for all A for which $\vdash_{\mathbf{KDH}} A$. But $\mathcal{M}^+$ is not a model for **KDH** because the 'new points' $\bar{w}$ in W^+ need not verify **H**, and any effort to fix this by ensuring that every point has a special successor then prevents the $\tau_{\Diamond\Box}$-version of Lemma 3 from going through.

Finally, returning to the theme of Section 1, the fact that $\mathbf{KX} \supsetneq \mathbf{K}$ does not prevent the translational embedding via $\tau_{\Box\Box}$ from succeeding even though, in the notation of note 7, $D_{\mathbf{KX}} \subsetneq D_{\mathbf{K}}$, since every distinction in **K** between formulas A and B is matched by a corresponding distinction in **KX** between $\tau_{\Box\Box}(A)$ and $\tau_{\Box\Box}(B)$.

[19] In calling a formula *valid at* a point in a frame, we mean that it is true at that point in any model on that frame.

Appendix

We begin by showing, as claimed in Section 3, that the canonical frame for any consistent normal modal logic extending **KH** satisfies the condition there called H, repeated below for convenience. The purpose of including this standard style of argument here is to provide a contrast with the more interesting completeness result given later (involving the the condition H^+).

$$\mathsf{H} \qquad \forall w, x, y((Rwx \wedge Rwy) \rightarrow \exists u(Rwu \wedge \forall v(Ruv \rightarrow (Rxv \wedge Ryv)))$$

Let $\langle W, R\rangle$ be the canonical frame of any consistent normal modal logic in which **H** is provable. We need only show that this frame satisfies the condition H. It has to be shown how, given $w, x, y \in W$, one may find a $u \in W$ behaving as required: a $u \in R(w)$ with $R(u) \subseteq R(x) \cap R(y)$. We construct such a u as any maximally consistent (w.r.t. the logic in question) superset of the following set, once the set's own consistency has been established; it is clear enough that this will guarantee that u behaves as required:

$$\{A \,|\, \Box A \in w\} \cup \{\Box B \,|\, \Box B \in x\} \cup \{\Box C \,|\, \Box C \in y\}.$$

By familiar steps, we get that if this set were not consistent then we should have ("$\vdash$" indicating provability in the logic in question):

$$\vdash (A_1 \wedge \ldots \wedge A_k) \rightarrow \neg(\Box B_1 \wedge \ldots \wedge \Box B_m \wedge \Box C_1 \wedge \ldots \wedge \Box B_n)$$

where the A's, B's and C's are drawn from the three terms of the union inset above; further routine steps follow:

$$\vdash (A_1 \wedge \ldots \wedge A_k) \rightarrow \neg(\Box B_1 \wedge \ldots \wedge \Box B_m \wedge \Box C_1 \wedge \ldots \wedge \Box C_n)$$

$$\vdash (A_1 \wedge \ldots \wedge A_k) \rightarrow \neg(\Box(B_1 \wedge \ldots \wedge B_m) \wedge \Box(C_1 \wedge \ldots \wedge C_n))$$

$$\vdash (A_1 \wedge \ldots \wedge A_k) \rightarrow ((\Diamond\neg(B_1 \wedge \ldots \wedge B_m) \vee \Diamond\neg(C_1 \wedge \ldots \wedge C_n))$$

$$\vdash (\Box A_1 \wedge \ldots \wedge \Box A_k) \rightarrow \Box(\Diamond\neg(B_1 \wedge \ldots \wedge B_m) \vee \Diamond\neg(C_1 \wedge \ldots \wedge C_n)),$$

at which point we invoke the an appropriate substitution instance of the axiom **H** to conclude that

$$\vdash (\Box A_1 \wedge \ldots \wedge \Box A_k) \rightarrow (\Box\Diamond\neg(B_1 \wedge \ldots \wedge B_m) \vee \Box\Diamond\neg(C_1 \wedge \ldots \wedge C_n)).$$

So, since each $\Box A_i \in w$ forces one or other disjunct of the consequent into w, thereby forcing either $\Diamond\neg(B_1 \wedge \ldots \wedge B_m)$ into $x \in R(w)$, contradicting the fact that each $\Box B_j \in x$, or $\Diamond\neg(C_1 \wedge \ldots \wedge C_m)$ into $y \in R(w)$, this time contradicting the provenance of the C's.

Another completeness result for **KH** was mentioned in Section 3, namely that this logic is determined by the class of frames satisfying:

$$\mathsf{H}^+ \qquad \forall w(R(w) \neq \varnothing \rightarrow \exists u \in R(w) \forall x \in R(w) \,.\, R(u) \subseteq R(x)).$$

We show this, too, by looking at the canonical frame $\langle W, R\rangle$ of any consistent normal extension of **KH**. (The condition is abbreviated here with the '$R(w)$' notation.)

We must show how to find, for any $w \in W$ with a successor, a special successor (to use Section 3's terminology): a u as promised by the consequent of H^+. So suppose $R(w) \neq \varnothing$ To obtain the desired $u \in W$, we take any maximal consistent extension of the set $\{A \mid \Box A \in w\} \cup \{\Box B \mid \Diamond\Box B \in w\}$ – leaving the reader to check that this suffices for u to serve as a special successor to w – assuring ourselves of the consistency of this set (relative to the logic in question) by the following reasoning. Its inconsistency would mean (with $\vdash$ as above) that for some A_i $(i = 1, \ldots, k)$ with $\Box A_i \in w$ and B_j $(j = 1, \ldots, n)$ with $\Diamond\Box B_j \in w$:

$$\vdash (A_1 \wedge \ldots \wedge A_k) \rightarrow \neg(\Box B_1 \wedge \ldots \wedge \Box B_n) \tag{1}$$

and so by the normality of $\Box$ and a little De Morgan on the consequent:

$$\vdash (\Box A_1 \wedge \ldots \wedge \Box A_k) \rightarrow \Box(\Diamond\neg\, B_1 \vee \ldots \vee \Diamond\neg B_n). \tag{2}$$

Since the antecedent here is in w, so is the consequent. But by $\mathbf{H}_n$, or more accurately, a variant of the dual of $\mathbf{H}_n$ as given in Section 3:

$$\vdash \Box(\Diamond\neg B_1 \vee \ldots \vee \Diamond\neg B_n) \rightarrow (\Box\Diamond\neg B_1 \vee \ldots \vee \Box\Diamond\neg B_n). \tag{3}$$

Thus one of the disjuncts $\Box\Diamond\neg B_j$ of the consequent must belong to w, contradicting the choice of the various B_j as formulas for which $\Diamond\Box B_j \in w$. What, however, of the possibility that there are no such B_j? This would mean that the negation of the antecedent in (1) was provable, which would contradict the fact that each $\Box A_i \in w$ and $R(w) \neq \varnothing$.

As a simple corollary, we have that the canonical frame for any consistent normal modal extension of **KDH** satisfies the condition DH^+ from Section 3 (H without its antecedent), and thus that **KDH** itself is determined by the class of frames meeting this condition. Also, we note that since whenever a class of frames contains the canonical frame for some logic but not every frame for the logic, that class is not modally definable, it follows that the class of frames satisfying H^+ is not modally definable.

There is a point of some theoretical interest here arising over the extent to which a condition (+), in general *stronger* than a corresponding condition (−), is *equivalent* to it when considered specifically as a condition on canonical frames for normal modal logics. (What follows is a tentative and provisional attempt to characterize the phenomenon of interest; a more general formulation may be desirable, and the use just made of the phrase "in general" turns out to fall short of meaning "universally". This will become evident presently, *à propos* of **KH**, and it suggests that these remarks will need to be corrected in other ways too.) We consider the first-order language with dyadic predicate letter R as sole non-logical primitive. Where $\Phi(w, x, u)$ is a formula of this language in which at most the exhibited variables are free – though possibly containing further quantifiers and their associated bound variables – let $\Phi^{\wedge}_n(w, x, u)$ be the result of replacing any subformula Rxv (v any variable distinct from x[20]) in $\Phi(w, x, u)$ by

[20] The author has not considered the case of subformulas Rxx, or, for that matter, $x = v$.

$Rx_1v \wedge \ldots \wedge Rx_nv$, with an analogous replacement in the case of Rvx. Then the $(-)$ and $(+)$ conditions induced by Φ are the following:

$(-)$ For all $n \in Nat$: $\forall w \forall x_1 \ldots \forall x_n \exists u \Phi_n^{\wedge}(w, x, u)$

$(+)$ $\forall w \exists u \forall x \Phi(w, x, u)$.

(In $(-)$, the $n = 0$ case is to be understood as for the modal object language: see note 17.) For the case of **KDH**, $(+)$ is DH^+, here rewritten into the above format (with the key quantifier prefix preceding a suitable Φ):

$$\forall w \exists u \forall x [Rwu \wedge (Rwx \to \forall y (Ruy \to Rxy))]$$

and the corresponding $(-)$ condition is the combined requirement that for each $n \in Nat$:

$$\forall w \forall x_1 \ldots \forall x_n \exists u [Rwu \wedge (\bigwedge_{i=1}^{n} Rwx_i \to \forall y (Ruy \to \bigwedge_{i=1}^{n} Rx_iy))].$$

If the canonical frame for a normal modal logic satisfies the latter – (in general) weaker – condition then the logic must prove $\mathbf{H}_n$ for all $n \in Nat$, which in turn guarantees that the frame satisfies the former – stronger – condition. (For a variation inspired by the case of **KH** the $(+)$ condition would be

$$\forall w \exists u \forall x [Rwx \to (Rwu \wedge \forall y (Ruy \to Rxy))],$$

which is equivalent to our condition H^+, which is a "canonical consequence" of the corresponding $(-)$ condition, requiring for all n:

$$\forall w \forall x_1 \ldots \forall x_n \exists u [\bigwedge_{i=1}^{n} Rwx_i \to (Rwu \wedge \forall y (Ruy \to \bigwedge_{i=1}^{n} Rx_iy))].$$

By contrast with the **KDH** weak condition, the $n \geq 1$ cases of this condition do not imply seriality, and the $n = 1$ case is a logical truth, matching the vacuity noted in Section 3 of $\mathbf{H}_1$; further, as in the discussion there, the m and n cases are equivalent whenever $m, n \geq 2$. This $(-)$ condition can accordingly be simplified to consist of the requirements given by the $n = 0$ and $n = 2$ cases, the latter being our condition H. But the former implies seriality, which, as noted, does not follow from the present $(+)$ condition, contrary to our stage-setting remarks characterizing the $(-)$ condition as weaker than – implied by – the stronger $(+)$ condition. This anomaly shows that the author's current grasp of the general situation leaves much to be desired.)

A simpler example, both in reducing the complexity of Φ and in not involving the anomaly just parenthetically alluded to, of the $(-)$ form implying the $(+)$ form over canonical frames is the following, which may already be familiar to the reader. Here the initial $\forall w$ is absent (or vacuous) and $\Phi(w, x, u)$ is just Rux. The $(-)$ condition then says that any finitely many points have a common R-predecessor while the (+) condition says that some point bears the relation R to every point.[21] While in general the former condition is weaker than the latter, restricting attention to canonical frames, the "$(-) \Rightarrow (+)$" implication holds, since if the canonical frame for (normal modal) **S** satisfies $(-)$, then **S** enjoys the Rule of Disjunction, as it is called in [17], p. 44, from which one may conclude that the canonical frame for **S** satisfies the (+) condition.

A case intermediate in complexity is provided by the following generalized form of **G** (from Section 3), namely the schema:

$\mathbf{G}_n \quad (\Diamond\Box p_1 \wedge \ldots \wedge \Diamond\Box p_n) \rightarrow \Box\Diamond(p_1 \wedge \ldots \wedge p_n).$

For a particular choice of n, the formula $\mathbf{G}_n$ – which differs from $\mathbf{H}_n$ in having "$\Box\Diamond$" in place of "$\Diamond\Box$" in the consequent – is valid at precisely those points any $n + 1$ successors of which themselves have a common successor; in speaking of the schema $\mathbf{G}_n$, we have in mind the simultaneous presence of all such formulas. Their provability is guaranteed for a normal modal logic whose canonical frame satisfies the $(-)$ condition we get by taking $\Phi(w, x, u)$ as $Rwx \rightarrow Rxu$, which in turn suffices for the canonical frame to satisfy the corresponding (+) condition, according to which

$$\forall w \exists u \forall x (Rwx \rightarrow Rxu).$$[22]

It would be interesting to know what the range of conditions Φ is for which this '$\exists$-fronting' $(-)$-to-$(+)$ transition goes through (as well as clarifying the situation described parenthetically as anomalous *à propos* of the **KH**-inspired (+) and $(-)$ conditions). The whole issue is presumably closely connected with other compactness-induced phenomena in modal logic such as Fine's modal saturation property and the Goldblatt–Thomason–van Benthem ultrafilter extension construction, but exactly how this connection goes is not clear to the author – though it may well have been explored in the literature of the subject somewhere.

Acknowledgments. I am grateful to Allen Hazen and to Bryn Humberstone for their comments on the material in Section 4 of this paper, and to Rohan French for providing the challenge in response to which it was written, as well as to several careful referees for their corrections.

Note added in press. After this paper was completed, my attention was drawn to [25], which provides an elegant and highly informative discussion of the embedding possibilities amongst various (not necessarily normal) modal logics, though without special attention to our weaker/stronger theme.

[21] Such formulations call to mind a second-order representation of the $(-)$ conditions, with monadic second-order variables ranging over finite sets. Instead of the above $\Phi_n^{\wedge}(w, x, u)$ formulas we would have $\Phi^{\forall}$, say, in which subformulas Rxv are replaced by $\forall x(Xx \rightarrow Rxv)$ – and analogously for Rvx – and the $(-)$ form is given by: $\forall w \forall X \exists u \Phi^{\forall}(w, x, u)$. This may be more suggestive for sorting out the present topic.

[22] To obtain u given w in the canonical frame, take $u \supseteq \{A \mid \Diamond\Box A \in w\}$. The consistency of this set follows from the fact that a normal modal logic proving all instances of $\mathbf{G}_n$ is closed under the rule: *From* $A_1 \vee \ldots \vee A_k$ *to* $\Box\Diamond A_1 \vee \ldots \vee \Box\Diamond A_k$. (In fact the admissibility of this rule for a given k is equivalent to the provability of the particular instance $\mathbf{G}_{k-1}$.)

BIBLIOGRAPHY

[1] Lennart Åqvist, 'On Dawson Models for Deontic Logic', *Logique et Analyse* **7** (1964), 14–21.

[2] Jean-Yves Béziau, 'Classical Negation Can Be Expressed by One of its Halves', *Logic Journal of the IGPL* **7** (1999), 145–151.

[3] Michael Byrd, 'Eventual Permanence', *Notre Dame Journal of Formal Logic* **21** (1980), 591–601.

[4] A. V. Chagrov and M. Zakharyaschev, 'Modal Companions of Intermediate Propositional Logics', *Studia Logica* **51** (1992), 49–82.

[5] A. V. Chagrov and M. Zakharyaschev, *Modal Logic*, Clarendon Press, Oxford 1997.

[6] B. F. Chellas, 'Modalities in Normal Systems Containing the **S5** Axiom', pp. 261–265 in C. Diamond and J. Teichman (eds.), *Intention and Intentionality: Essays in Honour of G. E. M. Anscombe*, Harvester Press, Brighton 1979.

[7] B. F. Chellas, *Modal Logic: An Introduction*, Cambridge University Press, Cambridge 1980.

[8] E. E. Dawson, 'A Model for Modal Logic', *Analysis* **19** (1959), 73–78.

[9] K. Došen, 'Modal Translations in Substructural Logics', *Journal of Philosophical Logic* **21** (1992), 283–336.

[10] G. E. Hughes and M. J. Cresswell, *A New Introduction to Modal Logic*, Routledge, London 1996.

[11] Lloyd Humberstone, 'Contra-Classical Logics', *Australasian Journal of Philosophy* **78** (2000), 437–474.

[12] Lloyd Humberstone, 'Béziau's Translation Paradox', *Theoria* **71** (2005), 138–181.

[13] Lloyd Humberstone, 'Logical Discrimination', pp. 207–228 in J.-Y. Béziau (ed.) *Logica Universalis: Towards a General Theory of Logic*, Birkhäuser, Basel 2005.

[14] Lloyd Humberstone, 'Modality', Chapter 20 (pp. 534–614) in Frank Jackson and Michael Smith (eds.), *The Oxford Handbook of Contemporary Philosophy*, Oxford University Press 2005.

[15] Lloyd Humberstone, 'Modal Logic for Other-World Agnostics: Neutrality and Halldén Incompleteness', forthcoming, *Journal of Philosophical Logic*.

[16] Charles F. Kielkopf, 'K1 as a Dawson Modelling of A. R. Anderson's Sense of "Ought"', *Notre Dame Journal of Formal Logic* **15** (1974), 402–410.

[17] E. J. Lemmon and D. S. Scott, *An Introduction to Modal Logic*, ed. K. Segerberg, American Philosophical Quarterly Monograph Series, Basil Blackwell, Oxford 1977. (Originally circulated 1966.)

[18] Wolfgang Lenzen, 'Epistemologische Betrachtungen zu [S4, S5]', *Erkenntnis* **14** (1979), 33–56.

[19] K. Matsumoto, 'Reduction Theorem in Lewis' Sentential Calculi', *Math. Japonicae* **3** (1955), 133–135.

[20] F. J. Pelletier, 'Six Problems in "Translational Equivalence"', *Logique et Analyse* **27** (1984), 423–434.

[21] F. J. Pelletier and A. Urquhart, 'Synonymous Logics', *Journal of Philosophical Logic* **32** (2003) 259–285.

[22] N. Rescher and A. Urquhart, *Temporal Logic*, Springer-Verlag, New York 1971.

[23] H. Sahlqvist, 'Completeness and Correspondence in the First and Second Order Semantics for Modal Logic', pp. 110–143 in S. Kanger (ed.), *Procs. Third Scandinavian Logic Symposium, Uppsala 1973*, North-Holland, Amsterdam 1975.

[24] K. Segerberg *An Essay in Classical Modal Logic*, Filosofiska Studier, Uppsala 1971.

[25] Evgeni E. Zolin, 'Embeddings of Propositional Monomodal Logics', *Logic Journal of the IGPL* **8** (2000), 861–882.

Lloyd Humberstone

Department of Philosophy, Monash University, Vic. 3800, Australia

Lloyd.Humberstone@arts.monash.edu.au