Synonymous Logics

Francis Jeffry Pelletier and Alasdair Urquhart

October 11, 2001

1 Introduction

Sometimes we have a feeling that two people or groups are "talking past one another." We feel, perhaps, that one is using a certain word A to mean X while the other uses A to mean Y, but this is compensated by the former using B to mean Y while the latter uses B to mean X. But since the people or groups don't recognize this – perhaps they *can't* recognize this – we see them as "doomed forever to misunderstand one another." Or, perhaps one group ("tribe") uses a certain linguistic construction to convey one type of information while another tribe uses it to convey a different type of information. But each tribe has some other way to convey the information which the opposed tribe uses this construction to convey. Again, the two tribes seem doomed to continual misunderstanding.

This situation need not be a case where just two words or constructions merely take on the role of one another. There may be many different words and constructions involved, and it could even be that there is no *one* word or construction of the one tribe's use which corresponds to some specific word or construction of the other tribe's use; it may instead be the cumulative interaction of many different words and constructions. In such a case trying to find a way of expressing one of these words as used by the second tribe back into the language of the first tribe would itself involve very many circumlocutions. The irony in these cases is that the two tribe *can* assert the same things, but they just don't realize it because they each mistakenly think the other group is using the words/constructions in the same way as they themselves do. And so they continue to "talk past one another."

The preceding discussion should bring to mind Quinean doctrines about "indeterminacy of translation" ([Quine, 1960]), especially the issue of whether there can be distinct translation manuals that are "empirically adequate" but nonetheless incompatible with one another. Examples of such cases are given in [Massey, 1978], where he discusses the possibility (in a first-order language with no singular terms) that the sets of individuals assigned to the predicates by one translation manual are the complements of the sets assigned by the other translation manual, the connectives and quantifiers in one manual are all assigned their logical duals in the other manual, and finally the speech acts of assertion and denial are opposite in the two manuals. (If uttering a sentence counts as asserting it according to one manual, then it counts as denying it according to the other manual.) Critics (e.g., [Kirk, 1982]) have objected to this last feature of Massey's method on the grounds that such speech acts should not be allowed as part of a translation manual. Although we ourselves take no stand on that issue, it is nevertheless an interesting issue whether there can be incompatible translation manuals that are "empirically adequate" but which do not differ from one another on any "speech act indicators." Our conclusion below is that there are such cases.

We wish to outline some conditions under which we can know this is happening; and, conversely, some conditions under which it can be proved that it is not happening. Since the same thoughts *can* be expressed in both languages – even if the speakers of the languages don't or can't recognize this due to their incompatible use of the subparts (the words or the constructions) of the languages – we might say that the languages are *equivalent* or *synonymous*. So our ultimate goal is to discover the conditions under which two languages are equivalent or synonymous.

We will investigate the specific version of this question as it arises with regard to different logics. In the above example, which considered natural language "misunderstandings," we imagined that the people involved had specific meanings associated with particular sentences and linguistic constructions. That is, they were communicating in an *interpreted* language. As we will be using the term, a *logic* is an *uninterpreted* set of symbols and rules for combining simpler uninterpreted components into larger uninterpreted components. Suppose we are given two logics for which we are interested in discovering whether they are "saying the same thing." From the fact that these logics are uninterpreted it follows that the main test will involve whether sentences are or are not theorems of their respective logics.

Although we will not directly be investigating the most general question of how to tell whether speakers of two interpreted languages are "talking past one another", our investigations both are interesting in their own right and could be extended to the more general case of interpreted languages if the primitive identities of meaning were known. One way this might happen is if it were known that such-and-so individual words in language A meant this-and-that in language B (and conversely) and it were also known how each linguistic construction was expressed in the other language. (It should be appreciated that merely knowing these things does not answer the question of whether the two languages give rise to "speaking past each other." There are many further conditions that would need to be met.) Alternatively, if, in each of the interpreted languages, the meanings of all the relevant items (words, constructions) were stated as axioms of that language, then our methods could be applied to discover whether the two languages were or were not capable of "saying the same thing."

So, we shall investigate the question of how to tell whether two logics – expressed as sets of symbols, axioms, and rules of inference – are or are not "really" the same logic, despite their differing symbols, axioms, and rules. We are all familiar with some simple examples of this phenomenon – for example, different formulations of classical propositional logic. However, we might ask ourselves *why* it is that we believe these to be "really the same logic" while at the same time believing that some other languages (e.g., the modal logics **S4** and **S5**) are *not* "really the same logic."

This paper is a sequel to an earlier paper by the first author [Pelletier, 1984b], in which the topic was called the problem of determining "translational equivalence" between logics, and six open problems are stated. The present paper answers these earlier problems by giving some general criteria for deciding whether or not two logics are translationally equivalent.

The basic problems can be listed as three main questions. The first, discussed in Section 2, asks: what is the correct definition of "translational equivalence" and "synonymy" between logics? The second, discussed in Sections 3 and 4, asks: are there simple, workable criteria for deciding whether two logics are equivalent, and therefore synonymous? And the third, discussed in Section 4, asks: what is the relation between mutual interpretability and translational equivalence?

In the section immediately following, we give precise definitions of the technical terms above, and in subsequent sections, give at least partial answers to these questions.

2 Translations between logics

Let us suppose that we are given two languages L_1 and L_2 for propositional logic. We shall assume that both languages contain an infinite supply of propositional variables $P_0, P_1, \ldots, P_n, \ldots$, together with a certain number of finitary connectives. As an example of such languages, we might consider the language of propositional modal logic, with a functionally complete set of classical connectives, and some modal operators (e.g., operators expressing possibility or necessity).

The logics that we consider will be given by axiom schemes and schematic rules; thus all logics will be closed under the rule of uniform substitution for propositional variables. If S is a logic, and A an axiom scheme that may involve connectives not in the language of S, then by $S + A$ we mean the logic in the language of S, together with any new connectives in A, containing the original axioms and rules of S, interpreted as applying to the expanded language, together with the axiom scheme A.

We wish to consider various examples of translation from one logic into another. As examples of such translations, we might consider Arthur Prior's translation of modal logic into tense logic; this is based on the Diodorean definition of "necessarily, A" as "it is now and always will be the case that A" [Prior, 1967]. Another familiar case is the double negation translation of classical logic into intuitionistic logic (see Example 2.3 below).

The notion of translation that we define in the next paragraphs is essentially the same as the concept of *schematic interpretation* of Prawitz and Malmnäs [Prawitz and Malmnäs, 1968]. A similar definition (though given in the context of logics defined by consequence relations) is given by Ryszard Wójcicki [Wójcicki, 1988, p. 69].

Let us add to the languages a new font P of variables $\alpha, \beta, \gamma, \delta, \ldots$, and call these new variables *parameters*. Let us call the languages we obtain from L_1 and L_2 by adding these new parameters $L_1(P)$ and $L_2(P)$. A *translation scheme* is an assignment from a subset of $L_1(P)$ to $L_2(P)$ taking the following form:

1. Every variable P_i in L_1 is assigned a formula A_i in L_2;

2. If f is a connective in L_1, then to the formula $f(\alpha_1, \ldots, \alpha_k)$ we assign a formula B of $L_2(P)$, where B contains only parameters from $\alpha_1, \ldots, \alpha_k$.

To illustrate this definition, here are four examples of translation schemes.

Example 2.1 *Let L_1 be the language of classical propositional logic based on the connectives $\wedge, \rightarrow, \leftrightarrow, \neg$, and L_2 the language of classical propositional logic based on $\vee, \neg$. Then we can express a familiar translation scheme from L_1 to L_2 as follows:*

1. *Each propositional variable is mapped into itself;*

2. $(\alpha \wedge \beta) \longmapsto \neg(\neg\alpha \vee \neg\beta)$;

3. $(\alpha \rightarrow \beta) \longmapsto (\neg\alpha \vee \beta)$;

4. $(\alpha \leftrightarrow \beta) \longmapsto \neg(\neg\alpha \vee \neg\beta) \vee \neg(\alpha \vee \beta)$;

5. $\neg\alpha \longmapsto \neg\alpha$.

So we see that L_1 $(\wedge, \rightarrow, \leftrightarrow, \neg)$ can't "say anything more" than can already be "said" in L_2 $(\vee, \neg)$. It is clear that there is an inverse translation from L_2 to L_2, using the familiar definition

$$(\alpha \vee \beta) \longmapsto \neg(\neg\alpha \wedge \neg\beta),$$

and that the two languages L_1 and L_2 are essentially equivalent. We shall see, however, in Section 4 below that in general having such two-way translations is insufficient for the translational equivalence of two systems.

Example 2.1 also illustrates a potentially significant feature of translation functions. The translation given is in fact *inefficient* in the sense that it may involve a potentially exponentially large increase in passing from a formula to its translation. This is because the translation of the biconditional $\leftrightarrow$ can double the size of a formula, and it is easy to find cases where the application of the translation function causes an exponential blowup in formula size.

In this paper, we shall ignore the question of the efficiency of translations, but it is a question of considerable significance in the area of the complexity of propositional proofs. It is possible to define an indirect translation function from L_1 to L_2 that avoids the exponential blowup. The reader is referred to [Krajíček, 1996] for the details.

Example 2.2 *Let L_1 be the language of intuitionistic propositional logic, and L_2 the language of the modal logic* **S4**. *Then Gödel's translation scheme from L_1 into L_2 is as follows:*

1. *Each propositional variable is mapped into itself;*
2. $\neg\alpha \longmapsto \neg\Box\alpha$*;*
3. $(\alpha \rightarrow \beta) \longmapsto (\Box\alpha \rightarrow \Box\beta)$*;*
4. $(\alpha \vee \beta) \longmapsto (\Box\alpha \vee \Box\beta)$*;*
5. $(\alpha \wedge \beta) \longmapsto \alpha \wedge \beta$*.*

Thus, there is a sense in which intuitionistic propositional logic cannot "say anything more" than **S4** "says."

Example 2.3 *Let L_1 be the language of classical propositional logic, and L_2 the language of intuitionistic propositional logic. Then the double negation translation of Kolmogorov [Troelstra and van Dalen, 1988, Vol. 1, p. 59] is as follows:*

1. *If P is a propositional variable, then $P \longmapsto \neg\neg P$;*
2. $(\alpha \wedge \beta) \longmapsto \neg\neg(\alpha \wedge \beta)$*;*
3. $(\alpha \vee \beta) \longmapsto \neg\neg(\alpha \vee \beta)$*;*
4. $\neg\alpha \longmapsto \neg\neg\neg\alpha$*.*

This time it is shown that there is a sense in which classical propositional logic cannot "say anything more" than intuitionistic propositional logic "says.".

Example 2.4 *Let L_1 be the language of pure implicational logic, and L_2 the language of propositional logic with disjunction and negation. Then we translate L_1 into L_2 as follows:*

1. *For a propositional variable P_i, $P_i \longmapsto P_{i+1}$;*

2. $(\alpha \vee \beta) \longmapsto ((\alpha \to \beta) \to \beta)$;

3. $\neg \alpha \longmapsto (\alpha \to P_0)$.

This last example is of interest as a translation scheme because it is not the identity mapping on propositional variables; the variable P_0 is singled out to play the role of the *falsum*. Translation schemes of this type are somewhat anomalous from the algebraic point of view. We shall say that a translation scheme is *simple* if it employs the identity map on propositional variables, and in addition, if the mapping

$$f(\alpha_1, \ldots, \alpha_k) \longmapsto A$$

is part of the translation scheme, then the formula A contains no variables other than the parameters $\alpha_1, \ldots, \alpha_k$.

If we are given a translation scheme between L_1 and L_2, then the translation determined by the scheme is the mapping $A \longmapsto A^t$ from formulas in L_1 to formulas in L_2 as given by the recursive definition:

1. If P_i is a variable in L_1, then $(P_i)^t = A_i$;

2. If f is a k-place connective of L_1, then

$$[f(A_1, \ldots, A_k)]^t = B[A_1^t/\alpha_1, \ldots, A_k^t/\alpha_k].$$

If S_1 and S_2 are proof systems in the languages L_1 and L_2, and t is a translation scheme from L_1 into L_2, then the translation scheme is *sound* if the translation A^t is provable in S_2 whenever A is provable in S_1, and *exact* if the translation A^t is provable in S_2 if and only if A is provable in S_1. A translation can be sound without being exact; as an example, consider the following translation:

Example 2.5 *Let t be the translation from the language of the modal logic* **K** *to the modal logic* **T** *that is the identity map on all formulas. Then t is sound, but not exact, since* **T** *has the theorem $\Box A \to A$, but* **K** *does not.*

The first four translations given above are all exact with respect to the appropriate systems. The first is an exact translation between two formulations of classical logic. The second is an exact translation from intuitionistic logic into **S4**. The third is an exact translation from classical into intuitionistic propositional logic, while the last is an exact translation from classical logic into its pure implicational fragment. We shall give further examples of exact translations in what follows.

Another very interesting translation in the literature is the translation from tense logic into modal logic in [Thomason, 1974, Thomason, 1975]. Thomason's translation method is very close to, but not identical with our notion of translation scheme, so we do not discuss it in detail here, but it is one of the most technically useful translations in the literature of modal logic.

Our main interest in the current paper is not with the fairly familiar notion of translation explained above, but rather with the stronger notion of translational equivalence introduced in [Pelletier, 1984b]. Let us suppose, as before, that we are dealing with two systems of propositional logic, S_1 and S_2, expressed in the languages L_1 and L_2. For reasons to be mentioned in the next paragraph, we shall assume that there is an equivalence connective $\leftrightarrow$ that is common to both systems. We assume for all the logics that we consider that they contain the following schematic axioms and inference rules governing this connective:

1. $\vdash A \leftrightarrow A$;
2. $A \leftrightarrow B \vdash B \leftrightarrow A$;
3. $A \leftrightarrow B, B \leftrightarrow C \vdash A \leftrightarrow C$;
4. $A_1 \leftrightarrow B_1, \ldots, A_k \leftrightarrow B_k \vdash O(A_1, \ldots, A_k) \leftrightarrow O(B_1, \ldots, B_k)$,

where O is any k-place connective in the system. The reason that we insist on these as explicit rather than derived rules is that we need to insure that they are preserved under extensions of the language in question. The conditions on the equivalence connective define all of the logics below as *equivalential logics* in the sense of [Prucnal and Wroński, 1974]. The model theory of equivalential logics is discussed in [Czelakowski, 1980].

Although in most of what follows, we shall be considering systems of modal logic based on classical logic, so that $\leftrightarrow$ can be taken to be the classical biconditional, it should be noted that the notion of translational equivalence given below is defined relative to whatever concept of equivalence connective that we single out in the logics, and there might be different equivalence connectives that could be reasonable choices for a given logic. For example, in a modal logic, we could equally well choose the classical biconditional or the strict biconditional, and these could lead to different results under the notion of translational equivalence. For this reason we insist that each logic in question have the same connective $\leftrightarrow$ in it which obeys the axiom schemes and rules given just above.

We say that S_1 and S_2 are *translationally equivalent* [1] if there are translation schemes t_1 and t_2 so that

1. Both t_1 and t_2 are sound;
2. For any formula in L_1, $(A^{t_1})^{t_2} \leftrightarrow A$ is a theorem of S_1;
3. For any formula in L_2, $(A^{t_2})^{t_1} \leftrightarrow A$ is a theorem of S_2.

[1] Pelletier in his original paper does not include the requirement that the equivalence is given in terms of translation schemes, but the examples and discussion in [Pelletier, 1984b] make it clear that this is in fact the notion intended.

Since we have assumed that the rule of detachment holds for the biconditional in the two systems, these conditions ensure that both t_1 and t_2 are exact. This can be seen as follows. If A^{t_1} is provable in S_2, then $(A^{t_1})^{t_2}$ is provable in S_1; but then, since $(A^{t_1})^{t_2} \leftrightarrow A$ is a theorem of S_1, A must also be a theorem of S_1.

An interesting example of translational equivalence between modal logics was given by the first author [Pelletier, 1984a]. In this case, both L_1 and L_2 are extensions of classical logic by modal operators $\triangle$ and $\Box$ respectively. The sentence $\triangle A$ is to be read as: "It is determinate that A"; in [Pelletier, 1984a], a number of axioms governing its use are given, and the resulting logic **ECNM∗** is proposed as a "logic of indeterminacy." The translation schemes for the two languages are then as follows:

1. $\triangle\alpha \longmapsto (\Box\alpha \vee \Box\neg\alpha)$;
2. $\Box\alpha \longmapsto (\triangle\alpha \wedge \alpha)$,

where it is understood that both translation schemes map variables and classical connectives into themselves. Pelletier shows that under these translation schemes, the logic **ECNM∗** is equivalent to the well-known modal logic **T**, and concludes that it "is in fact just logic **T** in disguise." A speaker of the one language would think s/he was talking about determinacy or maybe vagueness when using $\triangle$ and ∇, while speakers of the other language would think they were talking about necessity or perhaps possibility when using $\Box$ and $\Diamond$. But the fact that all their (logical) truths concerning the one concept can be expressed exactly and without residue by the other concept shows that neither group is entirely correct. They are actually talking about exactly the same thing, whatever it is. Either both groups are talking about both concepts–or they are talking about neither concept.

In a very interesting paper [Lenzen, 1979], Wolfgang Lenzen shows the translational equivalence of some well known systems of modal logic (details are given below). He uses a somewhat different concept of translational equivalence, which can be defined as follows. Let S be a system in a language L, and O a connective not in L. Then we say that a system S' in the language L' resulting from L by adding the new connective O to L is a *definitional extension* of S if it results from adding to S an axiom of the form

$$O(p_1, \ldots, p_k) \leftrightarrow A,$$

where A is a formula of L containing no variables other than $p_1, \ldots, p_k$. Being a definitional extension is a transitive relation: system S_3 is a definitional extension of a system S_1 if it is a definitional extension of a system S_2 that in turn is a definitional extension of S_1. We now say that two systems S_1 and S_2 are *synonymous* if there is a system S_3 in a language L_3 extending each of L_1 and L_2 so that S_3 is a definitional extension of both systems.

This concept of synonymy is the one employed in [Lenzen, 1979], and is essentially the same as the concept of "synonymous theories" employed in

[De Bouvère, 1965], in the context of model theory for classical logics. We now prove the equivalence to our own notion. This provides a partial answer to the first problem of [Pelletier, 1984b], which is as follows:

> **PROBLEM 1:** Is this notion of translational equivalence reasonable? That is, does it capture the intended force of "really the same system"? Is it in any sense trivial?

We shall assume in the next theorem that all translation schemes we consider are simple (this is necessary because of the restriction on variables in explicit definitions).

Theorem 2.6 *Two systems are synonymous if and only if they are translationally equivalent.*

Proof. To simplify notation, let us assume that S_1 and S_2 share a common language, except that $\Box_1$ and $\Box_2$ are unary connectives proper to L_1 and L_2.

($\Rightarrow$) Let S_1 and S_2 be synonymous, that is to say, there are explicit definition schemes $D1$:

$$\Box_1 B \leftrightarrow A_1(B)$$

and $D2$:

$$\Box_2 B \leftrightarrow A_2(B)$$

so that the systems $S_1 + D2$ and $S_2 + D1$ have the same theorems. We use S_3 to stand for either of these two equivalent systems. Let t_1 and t_2 be the corresponding translation schemes, that is to say, t_1 is the scheme:

$$\Box_1\alpha \longmapsto A_1(\alpha),$$

while t_2 is the scheme:

$$\Box_2\alpha \longmapsto A_2(\alpha).$$

We need to show that S_1 and S_2 are translationally equivalent under these translation schemes. As a preliminary to proving this, we note that for any formula A of L_1, $A \leftrightarrow A^{t_1}$ is a theorem of S_3. The proof is by induction on the complexity of the formula A. The analogous result holds for any formula of L_2.

Hence, if A is a theorem of S_1, then A^{t_1} is a theorem of S_3. Since S_3 is a conservative extension of S_2, it follows that $\vdash_{S_2} A^{t_1}$. The proof that t_2 is sound is exactly symmetrical.

We now prove the second condition in the definition of equivalence. We have $\vdash_{S_3} A \leftrightarrow A^{t_1}$ and $\vdash_{S_3} A^{t_1} \leftrightarrow (A^{t_1})^{t_2}$, so $\vdash_{S_3} A \leftrightarrow (A^{t_1})^{t_2}$, hence $A \leftrightarrow (A^{t_1})^{t_2}$ is provable in S_1 since S_3 is a conservative extension of S_1. The proof of the third condition is symmetrical, so the proof that the two systems are translationally equivalent is complete.

($\Leftarrow$) Assume that S_1 and S_2 are translationally equivalent under translationally equivalent functions t_1, t_2 as above. Let $D1$, $D2$ be the corresponding definitions. We need to show that $S_1 + D2$ and $S_2 + D1$ are deductively equivalent. If A is a theorem of S_1, then $\vdash_{S_2} A^{t_1}$, and since $A \leftrightarrow A^{t_1}$ is a theorem of $S_2 + D1$, A is also a theorem of $S_2 + D1$. Secondly, for any variable P, the equivalences

$$\Box_2 P \leftrightarrow ((\Box_2 P)^{t_2})^{t_1} \leftrightarrow (\Box_2 P)^{t_2},$$

are provable in $S_2 + D1$, so that $D2$ is a theorem of $S_2 + D1$. Hence, every theorem of $S_1 + D2$ is a theorem of $S_2 + D1$. The converse is proved symmetrically. □

As promised above, we now present Lenzen's example of translational equivalence. **S4.4** is the modal logic that results from **S4** by adding the axiom scheme $A \to (\Diamond\Box A \to \Box A)$. That is, **S4.4** is the logic **K** plus the axiom schemes

$$\Box A \to A$$

$$\Box A \to \Box\Box A$$

$$A \to (\Diamond\Box A \to \Box A)$$

This logic was first defined in [Sobociński, 1964], and was investigated in detail in [Zeman, 1971], where it is shown to be characterized by relational frames that are "almost **S5** frames" in the sense that they result from **S5** frames by adding a single world that is related to all the original worlds, but not conversely. **KD45** is the logic that results from **K** by adding the axiom schemes

$$\Box A \to \Diamond A$$

$$\Box A \to \Box\Box A$$

$$\Diamond\Box A \to \Box A$$

Then Lenzen's result can be stated as follows:

Example 2.7 *The systems* **S4.4** *and* **KD45** *are equivalent under the translation schemes:*

$$\Box_1 \alpha \longmapsto (\Box_2 \alpha \wedge \alpha)$$

$$\Box_2 \alpha \longmapsto \Diamond_1 \Box_1 \alpha,$$

where $\Box_1$ *is the necessity operator of* **S4.4** *and* $\Box_2$ *the necessity operator of* **KD45**.

The reader is referred to [Lenzen, 1979, Satz 3] for the proof that these schemes establish translational equivalence. Thus these two systems are in effect notational variants of each other. Lenzen gives a philosophical interpretation to this result. He interprets **KD45** as an epistemic/doxastic logic, by reading the operator $\Box_2 A$ as "a is convinced that A," where a is an ideal rational agent. With this reading, and using the translation given above, the system **S4.4** becomes the logic of true belief for an ideal agent. Lenzen also proves two other results on translational equivalence that show that **S4.3.2** and **S4.2** can be considered as the logic of knowledge for individuals who accept the equation 'knowledge = true belief' only for certain special instances. He establishes this by showing that **S4.3.2** and **S4.2** are equivalent to certain epistemic/doxastic logics, using the same translation scheme as above.

3 Translation Invariants

In the previous section, we showed that our concept of translational equivalence is reasonable in the sense that it corresponds exactly to another widely used notion of translational equivalence. In this section, we provide further evidence of the reasonableness of the concept in the sense that its model-theoretic consequences are what we should expect. From the semantic point of view, we would like to say that if two systems are translationally equivalent, then they differ only in their syntactic form, while their "structural properties" should be essentially the same. In this section, we prove that this holds true for some properties of models that we can consider as structural. Later, we shall use these structural invariants of logics as criteria to prove that certain logics are *not* translationally equivalent.

The first structural property that we consider is the number of models of a given cardinality. We need first to define what we mean by a model. Let S be a logic in a language L. A *matrix* $\mathcal{M} = \langle \mathcal{A}, D \rangle$ for L consists of an algebra $\mathcal{A}$ having an operator corresponding to each connective of L, and in addition, a non-empty subset D of the universe of $\mathcal{A}$, the set of *designated values* of the matrix $\mathcal{M}$. The cardinality of a matrix $\mathcal{M}$ is the cardinal number of its universe, which may be finite or infinite. A matrix is *reduced* if D contains exactly one element. A valuation V of L in $\mathcal{M}$ is an assignment of elements of $\mathcal{A}$ to all the variables in L; we extend the valuation to all of the formulas of L by the inductive definition:

1. $[\![P_i]\!] = V(P_i)$;
2. $[\![C(A_1, \ldots, A_k)]\!] = O([\![A_1]\!], \ldots, [\![A_k]\!])$,

where O is the operator in $\mathcal{A}$ corresponding to the connective C. We say that a matrix $\mathcal{M}$ is a *model* of S if, whenever A is a theorem of S, then for any valuation in $\mathcal{M}$, $[\![A]\!]$ is a designated value in $\mathcal{M}$. Let $\mathcal{M} = \langle \mathcal{A}, D \rangle$ and $\mathcal{N} = \langle \mathcal{B}, E \rangle$ be two matrices. We shall say that $\mathcal{M}$ and $\mathcal{N}$ are *isomorphic* if there is a mapping ϕ that is an isomorphism between the algebras $\mathcal{M}$ and $\mathcal{N}$, and in addition, $\phi(D) = E$.

We shall be particularly interested in certain types of model, those in which the biconditional operator in the algebra corresponds to equality. Let $\leftrightarrow$ be the operation in a matrix $\mathcal{M}$ that interprets the biconditional in a given system S. Then we say that $\mathcal{M}$ is an *algebraic matrix* if it satisfies the condition that for any elements a, b of the algebra $\mathcal{A}$ in $\mathcal{M}$, $(a \leftrightarrow b) \in D \Leftrightarrow a = b$. Thus an *algebraic model* of a system S is an algebraic matrix that is a model for S.

Let us say that a property is a *translation invariant* if whenever a system S_1 has the property, and S_2 is translationally equivalent to S_1, then S_2 also has the property. Before stating the next theorem, we introduce some model-theoretic concepts that are general versions of corresponding concepts in classical model theory introduced in [De Bouvère, 1965]. If $\mathcal{M}$ and $\mathcal{N}$ are matrices for languages L_1 and L_2, where L_1 is a sublanguage of L_2, then we say that $\mathcal{N}$ is a *definitional expansion* of $\mathcal{M}$ if it results from $\mathcal{M}$ by adding new operations to the algebra of $\mathcal{M}$ corresponding to the new connectives of L_2. If F is a class of matrices for a language L_1, and G a class of matrices for a language L_2, where L_1 is a sublanguage of L_2, then we say that the class G is a *definitional extension* of F if for every matrix $\mathcal{M}$ in F, there is a unique matrix $\mathcal{N}$ in G that is a definitional expansion of $\mathcal{M}$. We say that two classes of matrices F and G are *coalescent* if there is a class H of matrices that is a common definitional extension for both F and G.

Theorem 3.1 *If S_1 and S_2 are translationally equivalent, and M_1 and M_2 are the classes of their algebraic models, then M_1 and M_2 are coalescent.*

Proof. Let S_1 and S_2 be translationally equivalent under the translation schemes t_1 and t_2. By Theorem 2.6, there is a system S_3 that is a definitional extension of both S_1 and S_2. Let M_3 be the class of algebraic models of S_3. Then any model $\mathcal{M}$ of S_1 (S_2) can be expanded to a model of S_3 by adding operators whose definition is determined by their definitions in the system S_3. This model is uniquely determined by $\mathcal{M}$, since the definitional axioms for the new operators are theorems of S_3, and in any algebraic model validating these definitions, the new operations are uniquely determined by the original operations of $\mathcal{M}$. □

Corollary 3.2 *The number of non-isomorphic (reduced) algebraic models of given cardinality is a translation invariant.*

Proof. By Theorem 3.1, if S_1 and S_2 are translationally equivalent, then the classes M_1 and M_2 of their algebraic models are coalescent. It follows from this that there is a bijection between M_1 and M_2 and hence between the subsets of these classes of a given cardinality. □

From this corollary we can show that two logics are *not* synonymous if they have different numbers of models for some particular size of the domain of these models.

Another invariant that can be considered as a "purely structural" property of logics is the abstract algebraic structure of propositions in the theory. We can make this idea precise by introducing the notion of a canonical model. Let S be a logic in a language L. We define the *canonical matrix* for L as follows. The algebra $\mathcal{A}$ underlying the matrix has as its universe the set of all equivalence classes of formulas of L under the relation of provable equivalence, where we write the equivalence class of A as:

$$[A] = \{B : \vdash_S A \leftrightarrow B\}.$$

The operations on the algebra $\mathcal{A}$ are defined by:

$$O([A_1], \ldots, [A_k]) = [C(A_1, \ldots, A_k)],$$

where the operation O corresponds to the connective C. The fact that these operations are well defined follows from the axioms and rules assumed for the biconditional. Finally, we define the set of designated elements by:

$$D = \{[A] : \vdash_S A\}.$$

The next lemma is essentially the classical result of Lindenbaum.

Lemma 3.3 *The canonical matrix for a logic S is an algebraic model for S.*

Proof. Let V be a valuation in the canonical matrix for S. If A is a theorem of L, containing the variables $P_1, \ldots, P_k$, then each variable P_i in A is assigned an equivalence class $[B_i]$ of formulas in the canonical model. Then it follows by induction that $[\![A]\!] = [A(B_1/P_1, \ldots, B_k/P_k)]$. Since the theorems of S are closed under uniform substitution, it follows that A takes a designated value. The fact that the canonical matrix is algebraic follows from the definition. □

The canonical matrix for a logic is unique up to isomorphism, so that we are justified in talking about *the* canonical model for a fixed system.

Theorem 3.4 *Let S_1 and S_2 be translationally equivalent logics and S_3 a system that is a definitional extension of both S_1 and S_2. Then the canonical model of S_3 is a definitional expansion of the canonical models of both S_1 and S_2.*

Proof. Since S_3 is a conservative extension of S_i, $i = 1, 2$, it follows that the L_i matrix obtained from the canonical matrix for S_3 by deleting operations not corresponding to connectives in L_i is isomorphic to the canonical matrix for S_i. The theorem then follows from Theorem 3.1. □

Corollary 3.5 *Every logic is complete with respect to its algebraic models.*

Proof. It follows from the definition of the canonical matrix for a logic that a formula is a theorem of the logic if and only if it is valid in the canonical model. Since the canonical model is an algebraic model, the theorem follows. □

4 Equivalence of modal logics

In this section, we focus on the area of modal logics, in which we can prove specific results, both positive and negative. We shall assume that a modal logic is given in a language that is an extension of classical propositional calculus, with a full set of connectives such as $\wedge, \vee, \leftrightarrow, \rightarrow, \neg$ and so forth. We shall assume that all the logics we discuss contain the same set of the classical connectives, and that the translation schemes between logics preserve the classical connectives exactly. We assume throughout this section that the biconditional common to all of the logics is the classical biconditional. All the logics in this section will be obtained by adding additional non-classical connectives to the classical ones.

To simplify the discussion, we shall confine ourselves to one-place modal operators, though the discussion extends easily to multi-place connectives. All of the logics we consider will contain a complete set of axioms and rules for classical propositional logic. We describe such systems as *modal logics*.

We now define a class of models that will be central in the remainder of this section. We say that a *Boolean matrix* $\mathcal{M}$ is a matrix in which the algebra $\mathcal{A}$ is a Boolean algebra $\mathcal{B}$ together with a unary operator O defined on $\mathcal{B}$, and a filter $F \subseteq \mathcal{B}$, constituting the set of *designated* values in $\mathcal{B}$. We carry over the definitions of valuations and validity from earlier sections. It is understood that the Boolean connectives are given their usual interpretation in the Boolean algebra, while the modal operator is assigned values by the definition $[\![\Box A]\!] = O([\![A]\!])$.

Theorem 4.1 *If S is a modal logic, then A is a theorem of S if and only if it is valid in all reduced algebraic Boolean models of S.*

Proof. By Theorem 3.4, any modal logic is complete with respect to its algebraic models. The theorem follows directly from this result, because any algebraic model for a modal logic is a Boolean matrix. Finally, we observe that the canonical model for a modal logic is reduced, because if A is provable in such a logic, then $A \leftrightarrow \top$ is also provable, hence $[\![A]\!] = [\![\top]\!]$. □

In the following results, we make use of the fact that there is a tight link between a modal logic and its classical models. We shall assume that the logics we consider each contain only one modal operator; this restriction is not essential, but simply serves to simplify notation. In addition, we shall restrict the translation functions we consider. We shall assume henceforth without special mention that all translation functions (including those in the definition of translational equivalence) are simple in the sense defined in Section 2, and that they map the classical Boolean connectives onto themselves.

We define a modal logic to be *normal* if it contains the logic **K** and is closed under the rule of necessitation. That is to say, the logic contains the following axiom scheme and rule of inference:

1. $\vdash \Box(A \wedge B) \leftrightarrow (\Box A \wedge \Box B)$;

2. if $\vdash A$ then $\vdash \Box A$.

It is well known that Boolean matrices for normal logics can often be described in terms of relational frames. This follows from a representation theorem for Boolean algebras with operators. Let us say that a *modal algebra* is an algebra consisting of a Boolean algebra $\mathcal{B}$, together with an operator O on $\mathcal{B}$ satisfying the conditions:

1. $O(x \wedge y) = O(x) \wedge O(y)$;
2. $O(1) = 1$.

These two conditions are algebraic counterparts of the axiom scheme and rule of inference defining normal modal logics.

Let R be a binary relation on a finite set U. We can define a modal algebra on the Boolean algebra of all subsets of U by setting for $A \subseteq U$:

$$O(A) = \{x \mid \forall y(xRy \Rightarrow y \in A)\}.$$

That is to say, $O(A)$ is defined by the usual truth conditions for modal operators in relational frames. We say that the modal algebra defined in this way on all the subsets of U is the modal algebra $\mathcal{A}(R)$ determined by the relation R.

It can be shown [Jónsson and Tarski, 1951, Jónsson and Tarski, 1952] that any modal algebra is a sub-algebra of a modal algebra determined by a relation. The following weaker result is general enough for our purposes.

Lemma 4.2 *Any finite modal algebra is isomorphic to a modal algebra determined by a relation.*

Proof. Let $\mathcal{A} = \langle \mathcal{B}, O \rangle$ be a finite modal algebra. By Stone's representation theorem for Boolean algebras, $\mathcal{B}$ is isomorphic to the algebra of all subsets of a finite set U. Thus, we can assume that $\mathcal{B}$ is in fact an algebra of this type. For $x, y \in U$, define the relation R by:

$$xRy \Leftrightarrow x \notin O(U - \{y\}).$$

Then it is not hard to show that the operator O is identical with the operator defined from this relation R. $\Box$

If R is a relation on an underlying set U, and L a language of modal logic, then we can assign subsets of U to the variables in L, and then give values to the formulas of L relative to points in U in the usual way. It follows from our definitions that a formula A is valid in such a relational frame (that is to say, A is true at all points under all valuations) if and only if it is valid in the modal algebra $\mathcal{A}(R)$. If $\langle U, R \rangle$ is a finite relational frame, we shall say that it is a *model* for a logic S if $\mathcal{A}(R)$ is a model for S. The preceding lemma shows that from the point of view of validity there is no distinction between relational frames and algebraic matrices.

Theorem 4.3 *If S_1 and S_2 are two normal modal logics differing in the number of finite relational frames in a given cardinality validating them, then they are not translationally equivalent.*

Proof. Corollary 3.2 implies that S_1 and S_2 are not translationally equivalent if they have different numbers of reduced algebraic models in a given finite cardinality. Since Lemma 4.2 shows that there is a bijective correspondence between such reduced algebraic models and the corresponding finite relational models, the theorem follows. $\Box$

The preceding theorem, although of a simple character, allows us to settle the question of the translational equivalence for many of the better known modal logics at one blow.

Theorem 4.4 *Let S_1 and S_2 be normal modal logics, where S_1 is a proper sublogic of S_2, and both logics have the finite model property with respect to the relational semantics. Then S_1 and S_2 are not translationally equivalent.*

Proof. Since S_1 is a sublogic of S_2, every finite relational model for S_1 is a model of S_2. However, since both logics have the finite model property, there must be a finite relational frame validating S_1 that is not a model of S_2. Hence, the number of finite relational models for S_2 must be less than that for S_1, for some fixed cardinality. It follows from Theorem 4.3 that S_1 and S_2 are not translationally equivalent. $\Box$

Theorem 4.5 *The logics* **K**, **T**, **B**, **S4**, **S5** *are all distinct from the point of view of translational equivalence.*

Proof. The distinctness of all of these logics follows from Theorem 4.4, with the exception of the pair of logics **S4** and **B**. However, we can settle this case by using Theorem 4.3. **S4** is determined by the class of all finite reflexive, transitive relational frames, while **B** is determined by the class of all finite reflexive, symmetric relational frames. There are three distinct reflexive, transitive frames on a set of two elements, while there are only two reflexive, symmetric frames on the same set. The distinctness of **S4** and **B** follows by Theorem 4.3. $\Box$

The preceding theorems give answers answer to the third, fourth and fifth problems of [Pelletier, 1984b], which read as follows:

PROBLEM 3: Find translation functions f_1 and f_2 ... for pairs of the well-known modal logics.

PROBLEM 4: Formulate a criterion which will tell whether two arbitrary logics have such translation functions or not.

PROBLEM 5: Would modal logic really become easier if all systems were translationally equivalent to one another?

As is clear from what we have just proved, Problem 3 has a negative answer for all of the well-known modal logics, though some less well-known systems, such as those described in Section 2, are in fact translationally equivalent. The cardinality criterion formulated above provides us with an answer to Problem 4, which in spite of its simplicity, enables us to settle the question of equivalence for essentially all of the well-known systems.

Since translational equivalence for modal logics turns out to be somewhat rare, the condition for Problem 5 fails. However, there are certainly interesting one-way translations, such as the translation of tense logic into modal logic due to S.K. Thomason, and mentioned above in Section 2. This translation is technically very useful, since it allows the construction of certain "pathological" logics (such as a finitely axiomatized, undecidable normal modal logic) in a much easier way than if the usual formalism with one modal operator is employed.

The cardinality criterion also enables us to answer a natural question about the relation between one-way translations and translational equivalence. The question is whether having exact translations in both directions is sufficient for translational equivalence. We shall show by an example that this condition is in fact not sufficient. Similar examples are known in the context of classical logic. An example of this kind in the thesis [Montague, 1957] is reported in [De Bouvère, 1965].

The two logics constituting our example are formulated with a single modal operator $\Box A$, together with a propositional constant C. The first logic, $\mathbf{T}^*$, is the result of adding the constant C to $\mathbf{T}$, with no added special axioms. That is, $\mathbf{T}^*$ is $\mathbf{K}$ plus $(\Box A \to A)$ plus the constant C. The second logic, $\mathbf{KU}^*$, is the logic resulting from $\mathbf{K}$ by adding the axiom schemes $\Box C$ and $C \to (\Box A \to A)$. A relational frame for these logics consists of a relation R that interprets the modal operator in the usual way, together with a subset S of the universe U to interpret the constant C (that is to say, C is true at a point $x \in U$ if and only if x belongs to S).

Lemma 4.6 *The logic* $\mathbf{KU}^*$ *is complete with respect to relational frames satisfying the conditions:*

1. $xRy \Rightarrow y \in S$;
2. $x \in C \Leftrightarrow xRx$.

Proof. This is easily established by the usual canonical model construction. $\Box$

The logic $\mathbf{KU}^*$ is an extension of the logic $\mathbf{KU}$ that results from $\mathbf{K}$ by adding the axiom scheme $\Box(\Box A \to A)$ [Chellas, 1980, p. 140]; $\mathbf{KU}$ is complete with respect to the relational condition of weak reflexivity: $xRy \Rightarrow yRy$. We obtain $\mathbf{KU}^*$ from $\mathbf{KU}$ by adding the constant C that represents the reflexive worlds. To distinguish the two necessity operators, we write the modal operator of $\mathbf{KU}^*$ as $\Box_1$, and the modal operator of $\mathbf{T}^*$ as $\Box_2$.

Theorem 4.7

1. *There are exact translations t_1 and t_2 between the systems* $\mathbf{KU}^*$ *and* $\mathbf{T}^*$*:*

$$t_1 : \Box_1 \alpha \longmapsto \Box_2 (C \to \alpha);$$

$$t_2 : \Box_2 \alpha \longmapsto (\Box_1 \alpha \wedge \alpha),$$

   where t_1 and t_2 are the identity mapping on C;

2. $\mathbf{KU}^*$ *and* $\mathbf{T}^*$ *are not translationally equivalent.*

Proof. The second translation scheme t_2 from $\mathbf{T}^*$ to $\mathbf{KU}^*$ is clearly exact, since every relational model for $\mathbf{T}^*$ is also a relational model for $\mathbf{KU}^*$.

To show that $\mathbf{KU}^*$ is exactly embedded in $\mathbf{T}^*$ by the first translation scheme, assume that A is a non-theorem of $\mathbf{KU}^*$, so that A is refutable in a frame $\mathcal{F}$ satisfying the conditions of Lemma 4.6. Define a new frame $\mathcal{G}$ by making all the worlds in $\mathcal{F}$ reflexive. Let S be the relation on the original frame $\mathcal{F}$, and R the new relation on $\mathcal{G}$. Then we have: $xSy \leftrightarrow (xRy \wedge y \in S)$. We can hence prove that for any formula A of $\mathbf{KU}^*$, and any x in $\mathcal{F}$, that A is true at x if and only if A^{t_1} is true in $\mathcal{G}$. Thus, A^{t_1} is not a theorem of $\mathbf{T}^*$.

The finite frames validating $\mathbf{T}^*$ are finite reflexive relational frames, together with an arbitrary subset S of the frame. The finite relational frames validating $\mathbf{KU}^*$ consist of a weakly reflexive relation, and a subset S consisting of the reflexive elements in the frame. There are three structurally distinct reflexive relations on a two element set; adding the subset S, we find that there are ten structurally distinct relational frames on a two element set validating $\mathbf{T}^*$. However, there are only six structurally distinct quasi-reflexive relations on a two element set. Hence, the fact that the two systems are not translationally equivalent follows from Theorem 4.3. $\Box$

5 Concluding Remarks

Most of this paper is concerned with the problems enunciated in [Pelletier, 1984b] about "translational equivalence." The issue raised by [Pelletier, 1984b]'s Problem 1 was whether or not this notion was a reasonable account of "synonymous logics." Much of the present paper was devoted to this issue, since a negative answer would rob any of the further "Problems" of their interest. Our strategy was to consider a number of proposals in the literature that purport to show that two systems of logic are "equivalent to one another" in one way or another. For some of them we proved that they are the same notion as our "translational equivalence", while for some others we showed that where they differed from ours was precisely in placed that proved them to be inadequate as accounts of "synonymous logics." Our conclusion therefore was that "translational equivalence" indeed does give the intuitively correct account of "synonymous logics."

Having then a criterion for synonymous logics, we give a number of criteria that can be used to tell whether or not two arbitrary logics are in fact synonymous, and we apply these criteria to show that none of the well-known systems are synonymous. These results provided answers to [Pelletier, 1984b]'s Problems 3–5.

We started this paper with an allusion to Quinean "indeterminacy of translation." Various of our proofs and examples show that there can be distinct and indeed incompatible "meanings" assigned to some of the terms in a language by different "translation manuals," and yet all the "empirical evidence" that is even theoretically available to the linguist making such manuals could not distinguish between them. That this is so follows from the existence of translationally equivalent logics: the two translation functions are such that a sentence of one interpretation of the language could be translated into a sentence in a different interpretation of the language, but this differently-interpreted sentence can be translated back into a sentence of the initial interpretation which is logically equivalent to the initial sentence. And despite the fact that the two interpretations are distinct – and even incompatible with each other, as the linguist who is considering the two translation manuals will testify – none of the participants can tell this because of the fact that any conceivable way of distinguishing interpretation A of some particular sentence from interpretation B is counterbalanced by a reinterpretation of the terms used in making such a distinction.

In the sorts of examples we have considered, we find one translation manual interpreting '$\Box$' as meaning "it is necessarily true that" while the other one interprets it as "it is not vague that." Another example had '$\Box$' interpreted as "an ideal observer is convinced that" vs. as "it is a true belief that." Clearly these interpretations are distinct; yet, there is no way for anyone to tell which is in fact being employed because everything that is affirmed (= is a theorem) in one intepretation is also affirmed in the other, and the non-affirmed ("empirical") sentences of the one interpretation are always provably equivalent to some non-affirmed sentence of the other interpretation. There is just no way for any participant in these languages to say which interpretation they "really" are using. And unlike the method used in [Massey, 1978], we have not changed any aspects of the "speech acts" (such as what counts as affirming) between the two interpretations.

This seems as good an answer as can be given to the sixth problem identified in [Pelletier, 1984b]:

> **Problem 6:** How does all this [facts about translational equivalence] relate to Quinean indeterminacy of translation and the intertranslatibility of alternate conceptual schemes?

References

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

[Czelakowski, 1980] Czelakowski, J. (1980). Equivalential logics I, II. *Studia Logica*, 40:227–236, 335–372.

[De Bouvère, 1965] De Bouvère, K. (1965). Synonymous theories. In Addison, J., Henkin, L., and Tarski, A., editors, *Symposium on the Theory of Models*, pages 402– 406. North-Holland Publishing Company.

[Jónsson and Tarski, 1951] Jónsson, B. and Tarski, A. (1951). Boolean algebras with operators. Part I. *Amer. J. Math.*, 73:891–939.

[Jónsson and Tarski, 1952] Jónsson, B. and Tarski, A. (1952). Boolean algebras with operators. Part II. *Amer. J. Math.*, 74:127–162.

[Kirk, 1982] Kirk, R. (1982). On three alleged rivals to homophonic translation. *Philosophical Studies*, 42:409–418.

[Krajíček, 1996] Krajíček, J. (1996). *Bounded Arithmetic, Propositional Logic and Complexity Theory.* Cambridge University Press.

[Lenzen, 1979] Lenzen, W. (1979). Epistemologische Betrachtungen zu [S4,S5]. *Erkenntnis*, 14:33–56.

[Massey, 1978] Massey, G. (1978). Indeterminacy, inscrutability, and ontological relativity. In *Studies in ontology.* Basil Blackwell. American Philosophical Quarterly monograph series No. 9.

[Montague, 1957] Montague, R. M. (1957). *Contributions to the axiomatic foundations of set theory.* PhD thesis, University of California, Berkeley.

[Pelletier, 1984a] Pelletier, F. J. (1984a). The not-so-strange modal logic of indeterminacy. *Logique et Analyse*, 108:415–422.

[Pelletier, 1984b] Pelletier, F. J. (1984b). Six problems in "translational equivalence". *Logique et Analyse*, 108:423–434.

[Prawitz and Malmnäs, 1968] Prawitz, D. and Malmnäs (1968). A survey of some connections between classical, intuitionistic and minimal logic. In Schmidt, H., Schütte, K., and Thiele, H.-J., editors, *Contributions to Mathematical Logic*, pages 215–229. North-Holland.

[Prior, 1967] Prior, A. (1967). *Past, Present and Future.* Oxford Univerity Press.

[Prucnal and Wroński, 1974] Prucnal, T. and Wroński, A. (1974). An algebraic characterization of the notion of structural completeness. *Bulletin of the Section of Logic, Polish Academy of Sciences*, 3:45–52.

[Quine, 1960] Quine, W. (1960). *Word and Object.* MIT Press.

[Sobociński, 1964] Sobociński, B. (1964). Modal system S4.4. *Notre Dame Journal of Formal Logic*, 5:305–312.

[Thomason, 1974] Thomason, S. (1974). Reduction of tense logic to modal logic. I. *Journal of Symbolic Logic*, 39:549–551.

[Thomason, 1975] Thomason, S. (1975). Reduction of tense logic to modal logic II. *Theoria*, 3:154–169.

[Troelstra and van Dalen, 1988] Troelstra, A. S. and van Dalen, D. (1988). *Constructivism in Mathematics: An Introduction.* North-Holland, Amsterdam.

[Wójcicki, 1988] Wójcicki, R. (1988). *Theory of Logical Calculi.* Kluwer.

[Zeman, 1971] Zeman, J. J. (1971). A study of some systems in the neighbourhood of S4.4. *Notre Dame Journal of Formal Logic*, 12:341–357.

University of Alberta
Edmonton, Alberta

University of Toronto
Toronto, Ontario