LOGIQUE ET ANALYSE

NOUVELLE SÉRIE

27e Année Décembre 1984

108

SOMMAIRE

Publication trimestrielle du Centre National Belge de Recherches de Logique

NAUWELAERTS PRINTING S.A.
B 3000 LOUVAIN (Belgium)
Valkerijgang 24/30

SIX PROBLEMS IN "TRANSLATIONAL EQUIVALENCE"

by
Francis Jeffry PELLETIER

I wish to first formulate a criterion according to which two systems of logic might be said to be "really the same system" in spite of their having different vocabulary – especially where the difference is in the logical operators each has. One way of doing this might be to show that the theorems of the two systems are validated by precisely the same structures; but I am here interested in a more "syntactic" test for this. On an intuitive level, my desires could be expressed by saying that I want to be able to *translate* one system into the other, preserving theoremhood. For this reason I call it "translational equivalence" between the two systems.

It is often said that a standard formulation P_1 of propositional logic using $\{\neg;\&\}$ as connectives (plus some rule of inference) and a formulation P_2 using $\{\neg;\rightarrow\}$ (plus perhaps Modus Ponens) are "merely notational variants of one another", or "are the same logic", or (in my terminology) are "translationally equivalent." Why is this? Isn't it because the two translations

1. $(p \& q) \;=\; \neg(p \rightarrow \neg q)$
2. $(p \rightarrow q) \;=\; \neg(p \& \neg q)$

preserve all the theorems of the two logics, and furthermore, 3 and 4 are true?

3 The rule of inference chosen for '&' is, when the '&'s are replaced by '→' in accordance with 1, a derivable rule in P_2.
4. The rule of inference chosen for '→' is, when the '→'s are replaced by '&'s in accordance with 2, a derivable rule in P_1.

At least this is what is said in many textbook discussions. In theory, though, one should say something further; namely that the translations do not introduce any disturbance into the systems. For, so far, all that has been said is that the translations have to map theorems onto theorems. But we would also like to see that non-theorems get mapped onto statements that "say the same thing in the other language" preserve equivalance; that is, take a non-theorem, translate it into the other language and then translate the result back into the first language. The result of this should be equivalent to the original formula.

Let's make this a bit more precise for our logics P_1 and P_2. We say that P_1 and P_2 are translationally equivalent just because there are two functions f_1 and f_2 which will accomplish the required translations from P_1 to P_2 and from P_2 to P_1, respectively. f_1 is function from the language P_1 to P_2 and f_2 from P_2 to P_1. So f_1 has the effect of saying "Given P_2 I can express the language P_1", and f_2 says "Given the language P_1 I can express P_2". They are defined as follows:

$f_1(A) = A$, if A is atomic	$f_2(A) = A$, if A is atomic
$f_1(\neg A) = \neg f_1(A)$	$f_2(\neg A) = \neg f_2(A)$
$f_1((A \& B)) = \neg(f_1(A) \rightarrow \neg f_1(B))$	$f_2(A \rightarrow B) = \neg(f_2(A) \& \neg f_2(B))$

Now to guarantee that P_1 and P_2 are in fact translationally equivalent, the translation functions must meet certain criteria, to wit

A. If A is an axiom of P_1 the $f_1(A)$ is a theorem of P_2
B. If A is an axiom of P_2 the $f_2(A)$ is a theorem of P_1
C. If R is a rule of inference in P_1 then f_1 applied to each premise of R and to the conclusion of R must be a derivable rule of P_2
D. If R is a rule of inference in P_2 then f_2 applied to each premise of R and to the conclusion of R must be a derivable rule of P_1
E. $[f_1(f_2(A)) \leftrightarrow A]$ is a theorem of P_2
F. $[f_2(f_1(A)) \leftrightarrow A]$ is a theorem of P_1

This is to say, after translation from one system to the other, the original axioms are still provable, the rule of inference are derivable, and the translation functions themselves introduce nothing new.

More generally, two systems of logic S_1 and S_2 are said to be translationally equivalent if there are translation functions f_1 and f_2 which obey A through F.

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 trival?

In fact my interests concerning translational equivalence have to do with modal logics. I am interested in determining whether certain well-known modal logics are translationally equivalent to one another. For this purpose I shall consider only translation functions that map non-modal vocabulary into themselves. And I shall assume that each modal logic has the same set of atomic propositions, the same propositional connectives $\{\neg, \&, \vee, \rightarrow, \leftrightarrow\}$, the same propositional rules of inference, and the same class of modal-free wffs. In fact, the only difference to be encountered is that system i has the modal operators $\Box_i$ and $\Diamond_i$ where system k has $\Box_k$ and $\Diamond_k$. So the translation functions are the identity function on everything except the modal operators. So the question amounts to: can we find systems i and k for which there are f_1 and f_2 obeying A through F?

It is often easy to find *one* of the required two functions which will obey the restrictions A and C. But finding a pair of functions so as to also satisfy B, D, E, and F is much more difficult, and in fact I can give only an "artificial" example of it (below).

I give first three examples of the case where the "translation" can be effected in one direction. Logic K is generated from the propositional logic together with these rules and axioms (the notation follows Chellas) (1)

[RE] if $\vdash (A \leftrightarrow B)$ then $\vdash (\Box_k A \leftrightarrow \Box_k B)$
[RN] if $\vdash A$ then $\vdash \Box_k A$
[Def◇] $\vdash \Diamond_k A \leftrightarrow \neg \Box_k \neg A$
[M] $\vdash \Box_k (A \& B) \rightarrow (\Box_k A \& \Box_k B)$
[C] $\vdash \Box_k (A \& \Box_k B) \rightarrow \Box_k (A \& B)$

System T is generated in the same way (with $\Box_t$ replacing $\Box_k$ together with the axiom

[T] $\vdash \Box_t A \rightarrow A$

Consider now f_1 (from T to K; that is, which says "given K, I can express T")

$$f_1(\Box_t A) = (\Box_k f_1(A) \& f_1(A))$$

It is easy to see that f_1 obeys A and C – that the translation of all T

(1) Brian CHELLAS *Modal Logic*, Cambridge U.P. 1980.

axioms are theorems of K, and the translations of the rules of T are derivable in K. For example,

$$\vdash \Box_t (A \& B) \rightarrow (\Box_t A \& \Box_t B)$$
$$\vdash \Box_t A \rightarrow A$$

are axioms of T whose translations are, respectively

$$\vdash (\Box_k (A \& B) \& (A \& B)) \rightarrow ((\Box_k A \& A) \& (\Box_k B \& B))$$
$$\vdash (\Box_k A \& A) \rightarrow A$$

One can easily see that, since [M] is already present in K, the first must be a theorem of K. And the second is propositionally a theorem of K. The translation of the rules, e.g., [RN], yields

if $\vdash A$ then $\vdash (\Box_k A \& A)$

which is obviously in K, since [RN] is in K.

Thus f_1 satisfies A and C, and so in some sense T is definable or modelable within K. The next task would be to find an f_2 such that f_2 $(\Box_k A) = X$, where X is some sentence of T. I am unable to find such a function which will obey B and D, and such that E and F are true.

There are various other examples, perhaps more interesting, of this kind of "one-way modelling". One of the deontic modal systems, KD45, is defined as K above (with $\Box_{d45}$ replacing $\Box_k$) plus these axioms

[D] $\vdash \Box_{d45} A \rightarrow \Diamond_{d45} A$
[4] $\vdash \Box_{d45} A \rightarrow \Box\Box_{d45} A$
[5] $\vdash \Diamond_{d45} A \rightarrow \Box \Diamond_{d45} A$

The system $S_{4.2}$, one of the systems often cited for use in tense logic, is the system T above (with $\Box_{4.2}$ replacing $\Box_t$, plus [4] above (with $\Box_{4.2}$ replacing $\Box_{d45}$), plus the "Geach formula"

[G] $\vdash \Diamond\Box_{4.2} A \rightarrow \Box \Diamond_{4.2} A$

Consider now the translation function f_1 (from KD45 to $S_{4.2}$, i.e. "given $S_{4.2}$ I can define KD45")

$$f_1 (\Box_{d45} A) = \Diamond\Box_{4.2} f_1 (A)$$

A quick check will yield the result that the translation of all KD45 axioms and rules of inference will be theorems and derivable in $S_{4.2}$. For example, [Def $\Diamond$] becomes

$$\vdash \Box \Diamond_{4.2} A \leftrightarrow \neg \Diamond \Box_{4.2} \neg A$$

[D] becomes

$$\vdash \Diamond \Box_{4.2} A \rightarrow \Box \Diamond_{4.2} A$$

[4] becomes

$$\vdash \Diamond \Box_{4.2} A \rightarrow \Diamond \Box \Diamond \Box_{4.2} A$$

[5] becomes

$$\vdash \Box \Diamond_{4.2} A \rightarrow \Diamond \Box \Box \Diamond_{4.2} A$$

It is easy to check that all of these are theorems of $S_{4.2}$. The translation of the rules also is straightforward. [RN] of KD45 becomes

$$\text{if } \vdash A \text{ then } \vdash \Diamond \Box_{4.2} A$$

which is obviously derivable in $S_{4.2}$, since it has its own [RN] and [T]. [RE] of KD45 becomes

$$\text{if } \vdash (A \leftrightarrow B) \text{ then } \vdash (\Diamond \Box_{4.2} A \leftrightarrow \Diamond \Box_{4.2} B)$$

which is also derivable. (In the presence of propositional logic and [Def $\Diamond$], the rule

$$\text{if } \vdash (A \leftrightarrow B) \text{ then } \vdash (\Diamond A \leftrightarrow \Diamond B)$$

is equivalent to [RE]. So given $\vdash (A \leftrightarrow B)$, the rule [RE] of $S_{4.2}$ would yield $\vdash (\Box_{4.2} A \leftrightarrow \Box_{4.2} B)$, and then the modified [RE] would yield the desired $\vdash (\Diamond \Box_{4.2} A \leftrightarrow \Diamond \Box_{4.2} B)$.) Again I have not been able to find an f_2 which will translate $\Box_{4.2} A$ into some formula of KD45 while satisfying B, D, E and F.

The two examples just given (of K and T, and of KD45 and $S_{4.2}$) illustrate different points. System K is *included* in system T (yet here we find that all theorems of T can be interpreted as certain theorems of K). System KD45 and system $S_{4.2}$ are independent of one another and yet all the theorems of KD45 can be interpreted as certain theorems of $S_{4.2}$. For full coverage I should find an example of systems X and Y, where X is included in Y, and yet all theorems of X are theorems of Y. Well – the "change only the subscript of the box" translation will do here. Thus between K and T for example we would have

$$f_2(\Box_k A) = \Box_t f_2(A).$$

Since system K is included in system T, this translation obeys the restrictions of mapping axioms into theorems and primitive rules into derivable rules. Of course, this f_2 together with the earlier f_1 does *not* make K and T translationally equivalent in my sense because conditions E and F are not satisfied. Of the two,

$$\vdash f_2(f_1(A)) \leftrightarrow A$$

and

$$\vdash f_1(f_2(A)) \leftrightarrow A$$

only the first is correct. To see this, let A be $\Box_t p$. Then $f_1(A) = (\Box_k p \,\&\, p)$ and $f_2(f_1(A)) = (\Box_t p \,\&\, p)$. Sure enough, we have

$$\vdash_T (\Box_t p \,\&\, p) \leftrightarrow \Box_t p.$$

But now let A be $\Box_k p$. Then $f_2(A) = \Box_t p$, and $f_1(f_2(A)) = (\Box_k p \,\&\, p)$. But

$$(\Box_k p \,\&\, p) \leftrightarrow \Box_k p$$

is not a theorem of K. The f_1 and f_2 given here for systems K and T *do* satisfy conditions A through D, however. So, if X is a formula of T and Y is a formula of K, then we have

$$\vdash X \text{ iff } \vdash f_1(X)$$
$$\vdash Y \text{ iff } \vdash f_2(Y)$$

What is required by conditions E and F, and what is missing from these f_1 and f_2, is that: starting with a formula of the one system and translating it into some formula of the other system, we can get back to an *equivalent* formula of the original system by applying the other translation function. It is not just that theorems can be translated into theorems, but that the entire language can be translated back and forth without introducing anything new. I guess the moral of this last example is that mutual modeling is insufficient for translational equivalence. The translation functions themselves have to be "innocuous" – or perhaps one should say that if one of them introduces some non-innocuity, then the other must be able to conteract it.

PROBLEM 2: Does this "one-way modeling" have any significance at all? Does it, for instance, show that deontic logic is a part of tense logic (by modeling KD45 in $S_{4.2}$)? Does it *clarify* anything? For instance, if K is intuitively clear, is the given interpretation of T in K clarifying? Is the interpretation of K in T clarifying?

I now turn to my "artificial" example of two systems which I *can* show to be translationally equivalent. I say that it is an artificial example because one of the systems has not independently been discussed. I developed it in the context of trying to write a logic for vagueness([2]) – a logic wherein $\Box$ was to be interpreted as "it is definite that" and $\Diamond$ interpreted as "it is indefinite (vague) whether". I then considered certain intuitive principles about vagueness and tried to express them as theorems of the logic. For example, I wanted the usual interdefinability of $\Box$ and $\Diamond$. (I use a subscript 'v' for these operators).

$$\Diamond_v A \leftrightarrow \neg\Box_v \neg A$$

And I thought that a statement was vague if and only if it was not definite.

$$\Diamond_v A \leftrightarrow \neg\Box_v A$$

([2]) I first presented it at a Canadian Philosophical Association meeting in Halifax in 1981 as part of a commentary on John Heintz's "Might There be Vague Objects?" A greatly espanded version was read at a conference "Foundations of Logic" in Waterloo, Ontario in 1982. What is more or less the present version was presented at a Society for Exact Philosophy meeting in Athens, Georgia in 1984. In the course of this, I have had a chance to discuss it with many people, of whom I should especially mention Alistair Urquhart, Johan van Bentham, Michael Dunn, and Richard Routley. Routley has pointed me to a series of papers by himself and Montgomery in the late 1960's in *Logique et Analyse* wherein they are concerned to give a foundation for the usual modal systems in terms of a contingency operator ∇ and a non-contingency operator Δ, rather than necessity ($\Box$) and possibility ($\Diamond$) operators. If one changes their ∇ to $\Diamond$ and their Δ to $\Box$, and treats their axioms and rules as defining a new system (rather than a new definition of an old system), one gets the results described here.

This forces upon one the principle that if a statement is vague (definite) then its negation must be also

$$\Diamond_v A \leftrightarrow \Diamond_v \neg A$$
$$\Box_v A \leftrightarrow \Box_v \neg A$$

Certain other principles also seemed plausible to me, for example that if A and B were both definite, so must their conjunction be

$$\vdash (\Box_v A \,\&\, \Box_v B) \rightarrow \Box_v (A \,\&\, B)$$

If some formula was provable, then it was definite

$$\text{if} \vdash A \text{ then} \vdash \Box_v A$$

If two formulae were provably equivalent, then it should be provable that one was definite just in case the other was

$$\text{if} \vdash (A \leftrightarrow B) \text{ then} \vdash (\Box_v A \leftrightarrow \Box_v B)$$

Certain formulae are *not* true in this logic, for example

$$\Box_v A \rightarrow \Diamond_v A$$
$$\Box_v A \rightarrow A$$
$$\Box_v A \rightarrow \Box \Box_v A$$
$$\Diamond_v A \rightarrow \Box \Diamond_v A$$
$$\Box_v (A \,\&\, B) \rightarrow (\Box_v A \,\&\, \Box_v B)$$
$$\Box_v (A \rightarrow B) \rightarrow (\Box_v A \rightarrow \Box_v B)$$

These invalid formulae will be recognized respectively as the analogues of [D], [T], [4], [5], [M], and an alternate foundation of logic K. At least I took them to be invalid under the interpretation of $\Box_v$ as "definite". A false sentence might be definite, so we don't have [T]; I thought that borderline examples of definiteness ("it's definite, but not definitely so") showed against [4] and [5]; a conjuction might be definite but neither conjuct be (as when they contradict each other); and an implication might be definite, and its antecedent definite, without its consequent being definite (as for example $((p \,\&\, \neg p) \rightarrow p)$). Although the analogue of [M] is not valid, as closely related principle does seem to be: (3)

(3) I shall not argue for this here. A brief plausibility argument appears in my "The Not-So-Strange Modal Logic of Indetermacy", *Logique et Analyse 108, pp. 415-422.*

$$\vdash \Box_v(A \,\&\, B) \,\&\, (A \,\&\, B)) \rightarrow (\Box_v A \,\&\, \Box_v B),$$

a principle I call [M*] ("if a conjunction is definitely true, then each conjunct must be definite").

As it turns out, this logic can be axiomatized by the following, added on top of propositional logic:

[RE]	if $\vdash(A \leftrightarrow B)$ then $\vdash(\Box_v A \leftrightarrow \Box_v B)$
[Def◇]	$\vdash \Diamond_v A \leftrightarrow \neg\Box_v \neg A$
[RN]	if $\vdash A$ then $\vdash \Box_v A$
[C]	$\vdash(\Box_v A \,\&\, \Box_v B) \rightarrow \Box_v(A \,\&\, B)$
[M*]	$\vdash \Box_v(A \,\&\, B) \,\&\, (A \,\&\, B)) \rightarrow (\Box_v A \,\&\, \Box_v B)$
[V]	$\vdash \Box_v A \leftrightarrow \Box_v \neg A$

(The last axiom stands for "vagueness"). This is an "anti-normal" logic, since it neither includes nor is included in K – it is not included in K because it has principle [V] which K does not have, and it does not include K because it lacks principle [M] which K has. Using the Montague-Scott method, it is straightforward to give a "neighborhood semantics" for this logic. In the paper mentioned in footnote 3, I showed that it is complete for the class of "contrary", "partially supplemented" models that are "closed under intersections" and "contain the unit" (I ignore the details here except to remark that [RE] and [Def◇] are valid in any class of such models, that "contrary" models validate [V], that "partially supplemented" models get their name from the fact that "supplemented models" validate [M] and "partially supplemented" ones validate that subset for which [M*] holds, that "closed under intersections" models validate [C], and "models which contain the unit" validate [RN]. For further details see Chellas.)

Now I claim that this logic is translationally equivalent to T. Consider the translation functions

$$f_1(\Box_t A) = (\Box_v f_1(A) \,\&\, f_1(A))$$
$$f_2(\Box_v A) = (\Box_t f_2(A) \vee \Box_t \neg f_2(A))$$

First, let us look at the axioms and rules of T, and replace $\Box_t A$ by $f_1(\Box_t A)$ throughout. We get (note that by $f_1(\Diamond_t A)$ transforms into $(\Diamond_v A \vee A)$

[RE′] if $\vdash(A \leftrightarrow B)$ then $\vdash((\Box_v A \,\&\, A) \leftrightarrow (\Box_v B \,\&\, B))$

[Def◇′] $\vdash \neg(\Diamond_v \neg A \vee \neg A) \leftrightarrow (\Box_v A \& A)$
[C′] $\vdash ((\Box_v A \& A) \& (\Box_v B \& B) \rightarrow (\Box_v (A \& B) \& (A \& B))$
[M′] $\vdash (\Box_v (A \& B) \& (A \& B)) \rightarrow ((\Box_v A \& A) \& (\Box_v B \& B))$
[T′] $\vdash (\Box_v A \& A) \rightarrow A$

And finally, let us look at $f_1 (f_2 (A)) \leftrightarrow A$. In the only interesting case, where $A = \Box_v p$: here $f_2(\Box_v p) = (\Box_t p \vee \Box_t \neg p)$, and f_1 of this is $((\Box_v p \& p) \vee (\Box_v \neg p \& \neg p))$, hence we get

[V′] $\vdash ((\Box_v p \& p) \vee (\Box_v \neg p \& \neg p)) \leftrightarrow \Box_v p$

It can easily be verified that these are all theorems of system V.

Now let's do it the other way: take the axioms and rules of V and replace $\Box_v A$ in accordance with f_2 (note that $f_2(\Diamond_v A)$ becomes $(\Diamond_t A \& \Diamond_t \neg A)$).

[RE″] if $\vdash (A \leftrightarrow B)$ then $\vdash ((\Box_t A \vee \Box_t \neg A) \leftrightarrow (\Box_t B \vee \Box_t \neg B))$
[Def◇″] $\vdash ((\Diamond_t A \& \Diamond_t \neg A) \leftrightarrow \neg(\Box_t \neg A \vee \Box_t \neg\neg A))$
[RN″] if $\vdash A$ then $\vdash (\Box_t A \vee \Box_t \neg A)$
[C″] $\vdash ((\Box_t A \vee \Box_t \neg A) \& (\Box_t B \vee \Box_t \neg B)) \rightarrow (\Box_t (A \& B) \vee \Box_t \neg(A \& B))$
[M″] $\vdash ((\Box_t (A \& B) \vee \Box_t \neg(A \& B)) \& (A \& B)) \rightarrow ((\Box_t A \vee \Box_t \neg A) \& (\Box_t B \vee \Box_t \neg B))$
[B″] $\vdash (\Box_t A \vee \Box_t \neg A) \rightarrow (\Box_t \neg A \vee \Box_t \neg A)$

And now we replace into the translation functions, using $A = \Box_t p$. So, $f_1(A) = (\Box_v p \& p)$, and f_2 of this is $((\Box_t p \vee \Box_t \neg p) \& p)$, yielding

[T″] $\vdash ((\Box_t p \vee \Box_t \neg p) \& p) \leftrightarrow \Box_t p$

It is again easy to see that these are all theorems of T.

So V and T are translationally equivalent; and, I would claim, this makes them the same logic, just as our earlier formulations of the propositional logic are "really notational variants of each other". And this is so despite the fact that looking at them in the abstract would never allow one to determine this – after all, V was even an "anti normal" logic and its semantics could only be given via a neighborhood method while T can be given a relational semantics on possible worlds.

The fact that V and T appear to be so different on the surface and yet are translationally equivalent makes me wonder whether other logics might have this feature. For example, perhaps S_4 and T are translationally equivalent? Perhaps they are really the same logic? Maybe all modal logics (except S_5)(4) are really the same logic? Wouldn't that make modal logic easy?!

PROBLEM 3: Find translation functions f_1 and f_2 which obey restrictions A through F 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?

PROBLEM 6: How does all of this relate to Quinean indeterminacy of translation and the intertranslatibility of alternate conceptual schemes? For example, can a radical translator ever tell whether he is talking with a native speaker of V as opposed to a native speaker of T? (The native speaker of V has no single word for "necessity" but can assert as a theorem everything that the native speaker of T can assert as a theorem. The native speaker of T has no single word for "vagueness" but can assert as a

(4) S_5 has only a finite number of modal functions of one variable and hence no system in which there are an infinite number of distinct modalities (e.g., T) can be translationally equivalent to it, because there are only a finite number of "things that can be said" in S_5 using only the variable p whereas there are an infinite number of "things to be said" in T using only the variable p. All the other usual modal systems have an infinite number of modal functions of one variable. In passing it should also be remarked that the number of distinct modalities is no sure clue to whether two systems are translationally equivalent. For example, if one adds the [4] axiom to T, yielding S_4, the resulting system has 14 irreducible modalities. Yet it is translationally equivalent to the system gotten by adding the [4] axiom (with appropriate subscript) to V, and this system has only 6 irreducible modalties.

theorem everything that the native speaker of V can assert as a theorem.) (5)

University of Alberta
Dpt. of Philosophy
Edmonton
Canada T6G2E5

Francis Jeffry PELLETIER

(5) Thanks go to the referee for various helpful comments on an earlier draft of this paper. The work reported in this paper was supported in part by a Canadian NSERC grant (A5525), and in part by grant RO 245/12 from Deutsche Forschungs Gemeinschaft under the auspices of Christian Rohrer of Stuttgart. My thanks to Professors Rohrer and Franz Guenthner (of Tübingen) for their hospitality during the summer of 1983.