CDD: 511.3

TOWARDS A STRONGER NOTION OF TRANSLATION BETWEEN LOGICS[1]

MARCELO E. CONIGLIO

Group for Theoretical and Applied Logic (GTAL)
CLE and IFCH – State University of Campinas (UNICAMP)
P.O. Box 6133, 13083-970 Campinas, SP
BRAZIL

coniglio@cle.unicamp.br

Abstract: The concept of translation between logics was originally introduced in order to prove the consistency of a logic system in terms of the consistency of another logic system. The idea behind this is to interpret (or to encode) a logic into another one. In this survey we address the following question: Which logical properties a (strong) logic translation should preserve? Several approaches to the concept of translation between logics are discussed and analyzed.

Key-words: Translations between logics. Abstract logics. Combination of logics.

1. INTRODUCTION

The original notion of translations between logic systems can be traced back to the pioneering works from Kolmogorov (1925), Glivenko (1929), Lewis & Langford (1932), Gödel (1933) and Gentzen (1933). As it is well-known, translations between logics were introduced as a tool for proving the consistency of a logic system relative to the consistency of another system. In particular, Kolmogorov (1925) showed

[1]Financed by FAPESP (Brazil), Thematic Project ConsRel, grant 2004/14107-2.

that the use of the Law of the Excluded Middle, criticized by Brouwer, does not led to contradictions: the trick is to translate any subformula of a given formula by its double negation. This is the first example in the literature of a translation between logic systems (in this case, classical logic was translated into intuitionistic logic).

The first translations which appear in the literature, most of them being from classical logic into intuitionistic logic, were defined as being mappings $f : \mathcal{L}_1 \longrightarrow \mathcal{L}_2$ from (the set of formulas of) logic $\mathcal{L}_1$ into (the set of formulas of) $\mathcal{L}_2$ just satisfying the following property: If the formula φ is a theorem of logic $\mathcal{L}_1$ then the formula $f(\varphi)$ must be a theorem of logic $\mathcal{L}_2$. Using this property, plus other specific characteristics of the translation mappings and the logics, it can be proved that if the logic $\mathcal{L}_1$ is inconsistent, so is $\mathcal{L}_2$.

Later on, the concept of translation between logics was improved by requiring the following stronger condition: φ is a theorem of logic $\mathcal{L}_1$ if and only if the formula $f(\varphi)$ is a theorem of logic $\mathcal{L}_2$. Afterwards, other authors required another condition (which is the usual one accepted today), stating the following: if $\Gamma \vdash_{\mathcal{L}_1} \varphi$ (that is, if φ is derivable from the set of hypothesis Γ in logic $\mathcal{L}_1$) then $f[\Gamma] \vdash_{\mathcal{L}_2} f(\varphi)$ (that is, $f(\varphi)$ is derivable from the set of hypothesis $f[\Gamma]$ in logic $\mathcal{L}_2$).[2] When "then" is replaced by "if and only if" in the last condition, we have a *strong* or *conservative* translation. The expression "conservative" to refer to a mapping between logics with such characteristics is part of the folklore, and it was used for instance in Meseguer (1989), Bell (1988) and several publications of the Campinas group, GTAL (see, for instance, da Silva, D'Ottaviano & Sette (1999) and D'Ottaviano & Feitosa (2001)). Consult also Carnielli & D'Ottaviano (1997) for historical remarks on translations between logics.

Of course a general notion of translation between logics presupposes a general definition of what a logic system is. There are different approaches to general logic systems in the literature, some of which will be briefly reviewed in sections 2 and 3.

---

[2] As usual, if $f : A \longrightarrow B$ is a function and $X \subseteq A$ then $f[X]$ stands for the set $\{f(x) : x \in X\}$.

Once we are positioned in the realm of logics defined in a wide sense (and not restricted to specific logic systems), it is natural to think about a translation $f : \mathcal{L}_1 \longrightarrow \mathcal{L}_2$ as being a *morphism between logics* (seen as formal structures) which relates $\mathcal{L}_1$ with $\mathcal{L}_2$, preserving or transferring (some of) the features of the logic $\mathcal{L}_1$ into the logic $\mathcal{L}_2$. And of course, if f is a conservative translation, we are tempted to infer that $\mathcal{L}_1$ is "encoded" into logic $\mathcal{L}_2$: after all, if we assume (and, in fact, we will do that) a Tarskian perspective, then a logic system is nothing more than a set of formulas together with a relation (the consequence relation) between sets of formulas and formulas. Thus, the preservation of that relation by a conservative translation would reveal that, as structures, $\mathcal{L}_2$ "contains" $\mathcal{L}_1$. (Probably we should add the requirement that f is an injective or even a bijective mapping.) As we shall see in the following sections, the question is not so simple.

The problems we address in this survey are the following: what is *in fact* preserved by a translation between logics, and what *should be* preserved? What are the properties that a mapping between logics should have in order to faithfully translate a logic into another? In this article we defend the thesis that the analysis of these questions only is possible at the meta-level, that is, by means of a formal study of the meta-properties of the logics.

This article is organized as follows: in Section 2 and 3 we briefly review some of the approaches to a general theory of logic systems appearing in the literature. In Section 4 we address the problem of the preservation of certain kind of meta-properties of logics by translations, in the context of combinations of logic. Finally, in Section 5 we will give the final reflections on the question of translating faithfully a logic system into another.

2. LOGICS IN A GENERAL SETTING

As we mentioned in the Introduction, a general theory of translations between logics should rely on a theory of general logic systems. This section is devoted to survey some of the proposals of a general definition of logic system which appeared in the literature. By no ways this brief review is complete, but we consider that it is enough

for our purposes: to discuss in a broad perspective the problem of stating a general definition of translation between logics.

The first general definition of logic system is due to Tarski, in his famous article *On Some Fundamental Concepts of Metamathematics*, published in 1930, based on a lecture of him from 1928 (see Tarski (1956)). However, it should be stressed that Paul Hertz, in 1929, already introduced an abstract logic system satisfying analogous properties, but in terms of a consequence relation (cf. Hertz (1929)). Hertz's system inspired Gentzen to introduce his celebrated sequent calculi. See Béziau (forthcoming) for a survey on the historical development of Tarski's notion of logical consequence.

According to Tarski, a logic in a broader sense can be characterized by a set S of entities, called *sentences* (or *formulas*), together with a *consequence operator* Cn satisfying certain reasonable conditions. In formal terms:

Definition 2.1 (Tarski, 1930) Let S be a non-empty set of entities called *sentences*. A *logic system* is an ordered pair $\langle S, Cn\rangle$ such that Cn is a mapping from $\wp(S)$ to $\wp(S)$ (called a *consequence operator*), in symbols $Cn : \wp(S) \longrightarrow \wp(S)$, satisfying the following properties:[3]

Axiom 0 $\overline{\overline{S}} \leq \aleph_0$.

Axiom 1 If $X \subseteq S$ then $X \subseteq Cn(X)$.

Axiom 2 If $X \subseteq S$ then $Cn(Cn(X)) = Cn(X)$.

Axiom 3 If $X \subseteq S$ then

$$Cn(X) = \bigcup\{Cn(Y) \ : Y \subseteq X \text{ and } Y \text{ is finite}\}.$$

Axiom 4 There exists a sentence $x \in S$ such that

$$Cn(\{x\}) = S.$$

□

[3]As usual, the expression $\wp(S)$ will stand for the power set of S, that is, the set of all the subsets of S.

Axiom 0 states that the set of sentences (or formulas) must be denumerable, that is, with cardinal not greater than aleph zero. Axiom 1 states that from a given set of premises we can be derive any premise of that set. Axiom 2 states that from the set of consequences of a given set X we derive nothing more than the formulas already derived from the set X. Axiom 3 states that the logic must be *compact*, that is, if a formula is derivable from a set of premises then it must be derivable from a finite subset of it. Finally, Axiom 4 states that it must exist a *bottom particle*, that is, a trivializing sentence.

This ample definition of logic system encompasses a good part of the usual logic systems. It is worth noting that the following *monotonicity property* is derived from the axioms above:

Axiom 3' If $X \subseteq Y \subseteq S$ then $Cn(X) \subseteq Cn(Y)$.

However, several non-classical logics do not satisfy some of the axioms above. In particular, Axioms 0, 3 and 4 seems to be a bit strong, and today most authors introduce a general *Tarskian* logic system as being a pair $\langle S, Cn\rangle$ just satisfying Axioms 1, 2 and 3'.

Sometimes it is more convenient to introduce a logic system through a consequence relation instead of by means of a consequence operator. Thus, given a consequence operator Cn satisfying Axioms 1, 2 and 3', consider the relation $\vdash$ contained in $\wp(S) \times S$ such that, for every $X \subseteq S$ and every $x \in S$,

$$\text{(I)} \qquad X \vdash x \text{ iff } x \in Cn(X).$$

It is easy to see that $\vdash$ satisfies the following properties:

(r1) If $x \in X$ then $X \vdash x$.

(r2) If $X \vdash x$ and $X \subseteq Y$ then $Y \vdash x$.

(r3) If $X \vdash x$ and $Y \vdash X$ then $Y \vdash x$.[4]

On the other hand, any relation $\vdash \subseteq \wp(S) \times S$ originates, by means of the definition (I), a mapping $Cn : \wp(S) \longrightarrow \wp(S)$ satisfying Axioms 1, 2 and 3', provided that $\vdash$ enjoys properties (r1), (r2) and (r3). It fol-

lows immediately that there exist a bijective correspondence between consequence operators satisfying Axioms 1, 2 and 3' and consequence relations satisfying properties (r1), (r2) and (r3). As a matter of fact, it is easy to prove that any relation $\vdash$ satisfying (r1) and (r3) must satisfy (r2). Thus, a Tarskian logic system can be presented through a consequence operator satisfying Axioms 1, 2 and 3' or by a consequence relation satisfying properties (r1) and (r3).

Later on, it was incorporated another important ingredient to the definition of a logic system: the concept of abstract *logical connective*. Thus, the set S of sentences is considered to be generated by a set of operators (called *connectives*). In other words, the set S is an *abstract algebra*, as it was observed for the first time by A. Lindenbaum. Using this notion, J. Łoś and R. Suszko introduced in 1958 an additional axiom for a consequence operator (cf. Łoś & Suszko (1958)):

(R) If $X \subseteq S$ and $\varepsilon : S \longrightarrow S$ is a substitution[5] then

$$\varepsilon[Cn(X)] \subseteq Cn(\varepsilon[X]).$$

In terms of consequence relations, axiom (R) is equivalent to the following:

(R)' If $X \subseteq S$ and $\varepsilon : S \longrightarrow S$ is a substitution then

$$X \vdash x \text{ implies that } \varepsilon[X] \vdash \varepsilon(x).$$

This means that the logical inferences are stable by substitutions. Just for exemplifying, if $\{p_1, (p_1 \Rightarrow p_2)\} \vdash p_2$ is valid in a given logic (where $\Rightarrow$ denotes an implication) then we should expect that $\{\varphi, (\varphi \Rightarrow \psi)\} \vdash \psi$ is still valid, for every formula φ and ψ. A logic system enjoying property (R) (or, equivalently, (R)') is called *structural*.

In the seminal article *Abstract Logics* (cf. Brown & Suszko (1973)), it was initiated an extensive study of general logics, using an interest-

[4]The expression $Y \vdash X$ is an abbreviation for "$Y \vdash y$ for every $y \in X$".

ing analogy between logic and topology. In this article they introduce general logic systems called *closure spaces*, in analogy with Kuratowski's closure operator defined to characterize topological spaces. Formally:

Definition 2.2 (Brown & Suszko, 1973) Let S be a non-empty set of entities called *sentences*. A *closure space* is an ordered pair $\langle S, Cn\rangle$ such that $Cn : \wp(S) \longrightarrow \wp(S)$ is consequence operator satisfying Axioms 1, 2 and 3' defined above. □

Continuing with the analogy between abstract logic and general topology, the morphisms between closure spaces are defined as being closure-preserving mappings.

Definition 2.3 (Brown & Suszko, 1973) Consider two closure spaces $\langle S_1, Cn_1\rangle$ and $\langle S_2, Cn_2\rangle$. A *continuous mapping* f from $\langle S_1, Cn_1\rangle$ to $\langle S_2, Cn_2\rangle$, denoted $f : \langle S_1, Cn_1\rangle \longrightarrow \langle S_2, Cn_2\rangle$, is a mapping $f : S_1 \longrightarrow S_2$ such that $f[Cn_1(X)] \subseteq Cn_2(f[X])$ for every $X \subseteq S_1$. □

Note that this is exactly the definition of a continuous mapping between topological spaces given in general topology. In terms of consequence relations, this is equivalent to the following: if $\vdash_1$ and $\vdash_2$ denote the consequence relations corresponding to the given closure spaces, then, for every set $X \cup \{x\} \subseteq S_1$:

$$X \vdash_1 x \text{ implies that } f[X] \vdash_2 f(x).$$

We can say that this definition of translation inaugurated the modern era of translations between logic systems.

In the same article, the authors consider more sophisticated closure spaces in which the set S of sentences is an abstract algebra generated by a set of connectives. Consequently, it is introduced the following definition:

---

[5] A *substitution* is an endomorphism, that is, a mapping $\varepsilon : S \longrightarrow S$ such that $\varepsilon(c(x_1, \ldots, x_n)) = c(\varepsilon(x_1), \ldots, \varepsilon(x_n))$ for every n-ary connective c and every $x_1, \ldots, x_n \in S$.

Definition 2.4 (Brown and Suszko, 1973) A pair $\langle \mathcal{A}, Cn \rangle$ such that $\mathcal{A}$ is an abstract algebra with domain $|\mathcal{A}|$ and $\langle |\mathcal{A}|, Cn \rangle$ is a closure space is called an *abstract logic*. An abstract logic is said to be *structural* if it satisfies property (R) above. □

The definition of translation between abstract logics is obtained by generalizing in a natural way the concept of continuous mappings between closure spaces.

Definition 2.5 (Brown and Suszko, 1973) A *logical morphism* $f : \langle \mathcal{A}_1, Cn_1 \rangle \longrightarrow \langle \mathcal{A}_2, Cn_2 \rangle$ from an abstract logic $\langle \mathcal{A}_1, Cn_1 \rangle$ to an abstract logic $\langle \mathcal{A}_2, Cn_2 \rangle$ of the same type is a continuous mapping $f : \langle |\mathcal{A}_1|, Cn_1 \rangle \longrightarrow \langle |\mathcal{A}_2|, Cn_2 \rangle$ which is also an homomorphism $f : \mathcal{A}_1 \longrightarrow A_2$ of algebras.[6] □

Brown and Suszko's definition of logic morphism is what today is normally used as the definition of translation between logic systems (even in the context of combinations of logics). This definition of morphism between logics, whereas is well-behaved in the sense of Category Theory and it is easy to manipulate, it is however a bit tight for some "real" applications: most of the usual mappings between concrete logics are not logic morphisms.

A new concept of translation between abstract logics, more flexible than the notion of logical morphism, was introduced in Wójcicki (1988): a logic system is, again, a pair $\langle \Sigma, Cn \rangle$ such that $\Sigma = (\Sigma_n)_{n \in \mathbb{N}}$ is a *propositional signature*, that is, a family of sets such that Σ_n is the set of connectives of arity n.[7] The algebra of formulas is then

[6] An *homomorphism* $f : \mathcal{A}_1 \longrightarrow A_2$ of algebras of the same type τ is a mapping $f : |\mathcal{A}_1| \longrightarrow |A_2|$ such that, for every n-ary connective c of τ, $f(c^{\mathcal{A}_1}(x_1, \ldots, x_n)) = c^{\mathcal{A}_2}(f(x_1), \ldots, f(x_n))$ for every $x_1, \ldots, x_n \in |\mathcal{A}_1|$. Here, as usual, $c^{\mathcal{A}}$ denotes the interpretation in the algebra $\mathcal{A}$ of the n-ary connective c of τ.

[7] Of course, from the point of view of Universal Algebra, to give a signature Σ is equivalent to give a type τ; we prefer to slightly modify the original presentation of Wójcicki (1988) and introduce the concept of propositional signature, that will be used in the rest of this paper.

freely generated by Σ from a fixed denumerable set $\mathcal{V}$ of propositional variables (the *atomic formulas*). The set of formulas generated by Σ is denoted by $L(\Sigma)$. The consequence operator Cn is defined as in Brown & Suszko (1973), that is, satisfying Axioms 1, 2 and 3'. On the other hand, the translations are a bit more elastic:

Definition 2.6 (Wójcicki, 1988) Given two logics $\langle \Sigma, Cn\rangle$ and $\langle \Sigma', Cn'\rangle$, a *translation* $f : \langle \Sigma, Cn\rangle \longrightarrow \langle \Sigma', Cn'\rangle$ is a continuous mapping between the underlying closure spaces such that:

- There exists a formula $\gamma_0(p_1) \in \mathbb{L}(\Sigma')$ depending just on propositional variable p_1 such that $f(p) = \gamma_0(p)$ for every variable $p \in \mathcal{V}$.
- For every n-ary connective $c \in \Sigma_n$ there exists a formula

$$\varphi_c(p_1, \ldots, p_n) \in \mathbb{L}(\Sigma')$$

depending on the variables $p_1, \ldots, p_n$ such that, for every formulas $\psi_1, \ldots, \psi_n \in \mathbb{L}(\Sigma)$ it holds:

$$f(c(\psi_1, \ldots, \psi_n)) = \varphi_c(f(\psi_1), \ldots, f(\psi_n)).$$

□

It is worth noting that all the "classical" translations between concrete logics which appeared in the literature (such as the translations of Glivenko, Gödel and Gentzen) are, from the linguistic point of view, translations in this sense. This approach is more convenient because is more elastic, however it is more difficult to work with, and some categorial properties are lost.

The ideas of Brown and Suszko were generalized, in a certain sense, by J.-Y. Béziau in his research program on *Universal Logic*. Under his perspective, a logic system is just a pair $\langle S, \vdash\rangle$ such that S is a set of entities (sentences), not necessarily being a free algebra (that is, not necessarily generated by connectives), and $\vdash$ is a consequence relation without any required property. The concept of translation between such structures coincides with that of Brown and Suszko.

Thus, a translation $f : \langle S_1, \vdash_1 \rangle \longrightarrow \langle S_2, \vdash_2 \rangle$ between logics is a mapping $f : S_1 \longrightarrow S_2$ such that, for every set $X \cup \{x\} \subseteq S_1$:

$$X \vdash_1 x \text{ implies that } f[X] \vdash_2 f(x).$$

The main goal of universal logic is to determine the domain of validity of certain metatheorems (for instance, the completeness theorem), as well as to obtain general formulations of such metatheorems. See, for instance, Béziau (1994).

On the other hand, Campinas group on logic (GTAL) developed, during the second half of the 90's, a research program on translations between logics, with emphasis on conservative translations, in the same line as Brown and Suszko's closure spaces and their continuous maps. Despite some of the results and definitions obtained in this research were already presented in Brown & Suszko (1973), several interesting examples of conservative translations between "concrete" logics were given. See, for instance, da Silva, D'Ottaviano & Sette (1999), Feitosa (1997), D'Ottaviano & Feitosa (1999) and D'Ottaviano & Feitosa (2001).

Another important mark in the development of general logic systems was the introduction, in the article Goguen & Burstall (1984), of the concept of *institutions* (the interested reader would consult instead the paper Goguen & Burstall (1992), considered the basic paper on the subject). Inspired by institutions, the article Meseguer (1989) introduced the notion of *general logics* (see also Cerioli & Meseguer (1993)). Both concepts (institutions and general logics) heavily rely on Category Theory, so the reader not acquainted with the basic notions of Category Theory can skip the rest of this section.

Institutions generalize Tarski's notion of truth, by substituting "vocabularies" by (abstract) signatures, and "translations among vocabularies" by abstract (categorial) morphisms between objects, called signature morphisms. The set of sentences are then parameterized by abstract signatures (as it was done above, when the sets $L(\Sigma)$ of sentences were parameterized by "concrete" signatures Σ). In the general case, we substitute the constructor L by a functor $\mathsf{Sen} : \mathbf{Sign} \longrightarrow \mathbf{Set}$. Since Sen is a functor then, for every signature morphism $f : \Sigma \longrightarrow \Sigma'$,

there is a mapping $\mathsf{Sen}(f) : \mathsf{Sen}(\Sigma) \longrightarrow \mathsf{Sen}(\Sigma')$. On the other hand, a contravariant functor Mod assigns to every signature Σ its class of models, in such a way that, if $f : \Sigma \longrightarrow \Sigma'$ is a signature morphism then $\mathsf{Mod}(f) : \mathsf{Mod}(\Sigma') \longrightarrow \mathsf{Mod}(\Sigma)$ is a mapping between the respective classes of models (note that Mod is contravariant). This approach can be generalized a bit, by considering categories of models (instead of classes of models). In formal terms:

Definition 2.7 (Goguen and Burstall, 1984) An *institution* is a tuple $\mathcal{I} = \langle \mathbf{Sign}, \mathsf{Sen}, \mathsf{Mod}, \models \rangle$ such that

- **Sign** is a category (of *(abstract) signatures*);
- $\mathsf{Sen} : \mathbf{Sign} \longrightarrow \mathbf{Set}$ is a functor, assigning to every signature Σ a set $\mathsf{Sen}(\Sigma)$ of *sentences*;
- $\mathsf{Mod} : \mathbf{Sign} \longrightarrow \mathbf{Cat}$ is a contravariant functor, assigning to every signature Σ a category $\mathsf{Mod}(\Sigma)$ of *models*;
- $\models$ is a family indexed by the class $|\mathbf{Sign}|$ of signatures such that, for every signature Σ, $\models_\Sigma \subseteq |\mathsf{Mod}(\Sigma)| \times \mathsf{Sen}(\Sigma)$ is a relation such that, for every signature morphism $f : \Sigma \longrightarrow \Sigma'$ the following *satisfaction condition* holds:

  $$M' \models_{\Sigma'} \mathsf{Sen}(f)(\varphi) \quad \text{iff} \quad \mathsf{Mod}(f)(M') \models_\Sigma \varphi$$

  for every model $M' \in |\mathsf{Mod}(\Sigma')|$ and every sentence $\varphi \in \mathsf{Sen}(\Sigma)$. □

In the definition above, **Set** and **Cat** stand for the usual category of sets and the usual category of (small) categories, respectively. Also, $|\mathbf{C}|$ denotes, as usual, the class of objects of a category $\mathbf{C}$. The satisfaction condition expresses the invariance of truth under "change of notation". The theory of institutions also generalize the duality between theories and model classes occurring in (first-order) Model Theory. Institutions were introduced as an abstract model theory applied to Computer Science, appropriate for developing concepts of specification languages such as structuring of specifications, parameterization, implementation, refinement etc.

It is important to notice that a given institution characterizes a *single* logic system (presented "semantically"). This means that the multiplicity of "semantical consequence relations" contained in an institution constitute, in fact, just a unique consequence relation, expressed in different "vocabularies" or "notations". So, given an institution, a translation (a morphism in **Sign**) $f : \Sigma \longrightarrow \Sigma'$ is not anymore a basis for a logic translation, as in the case of Brown and Suszko's abstract logics, but represents just a "change of notation" for the unique logic system underlying the institution. In other words, the translations are not logic translations, but vocabulary translations. The correspondent notion of logic translation between institutions is given through the concept of morphisms of institutions.

Whereas institutions are very abstract logics presented "semantically", the *entailment systems* (cf. Meseguer (1989)) are a kind of (abstract) syntactical counterpart (consult also Cerioli & Meseguer (1993) for more details on general logics). Formally:

Definition 2.8 (Meseguer, 1989) An *entailment system* is a triple $\mathcal{E} = \langle \mathbf{Sign}, \mathsf{Sen}, \vdash \rangle$ such that

- **Sign** is a category (of *(abstract) signatures*);
- $\mathsf{Sen} : \mathbf{Sign} \longrightarrow \mathbf{Set}$ is a functor, assigning to every signature Σ a set $\mathsf{Sen}(\Sigma)$ of *sentences*;
- $\vdash$ is a family indexed by the class $|\mathbf{Sign}|$ of signatures such that, for every signature Σ, $\vdash_\Sigma \subseteq \wp(\mathsf{Sen}(\Sigma)) \times \mathsf{Sen}(\Sigma)$ is a consequence relation satisfying (r1) and (r3) such that, for every signature morphism $f : \Sigma \longrightarrow \Sigma'$ and every set of sentences $\Gamma \cup \{\varphi\} \subseteq \mathsf{Sen}(\Sigma)$, the following holds:

$$\Gamma \vdash_\Sigma \varphi \quad \text{implies} \quad \mathsf{Sen}(f)[\Gamma] \vdash_{\Sigma'} \mathsf{Sen}(f)(\varphi).$$

□

The same remark about the unicity of the consequence of a given institution applies to entailment systems: there is just one ("syntactical") consequence relation represented by a given entailment system.

On the other hand, each "notation system" (signature) determines an specific representation of the consequence relation on that signature.

As mentioned above, a logic translation between institutions can be obtained through the concept of morphisms of institutions. The definition of morphisms between institutions (as well as between entailment systems) is a bit complicated and goes out of the scope of this paper, therefore we will not give the details here. The basic idea is that they are truth-preserving translations from one logical system into another. This means that they generalize the notion of translation between abstract logics of Brown and Suszko. Moreover, it is also considered the concept of *conservative* morphisms, which generalize the idea of conservative translations.

This section was devoted to briefly review relevant points in the development of a theory of general logic systems. In the next section we will discuss a different approach to logic systems, based on Model Theory.

3. FORMAL METALANGUAGES FOR ABSTRACT LOGICS

When used in particular cases with limited scope, the meaning of a translation between logics is clear, as the historical examples in the literature show. However, when considered in the broader perspective of abstract logics, the question "What does a translation really preserve?" is in fact intriguing. If a logic $\mathfrak{L}_1$ is translated "in the best possible way" into a logic $\mathfrak{L}_2$, are there any distinctions between them?

The fundamental point here is the following: in order to give a satisfactory answer to these questions, it is necessary to develop a theory capable of expressing the general meta-properties of abstract logics, as well as explaining the full meaning of translations between them.

In the paper Coniglio & Carnielli (2002) it was introduced a formal framework based on Model Theory for representing, at the meta-level, abstract logics and their translations. All the definitions, results and examples contained in this section are taken from the above mentioned paper. This section presupposes some basic knowledge of Model Theory.

We could be tempted to claim that, if f is a conservative translation between logics $\mathfrak{L}_1$ and $\mathfrak{L}_2$, then the *meaning* of $\mathfrak{L}_1$ is preserved within $\mathfrak{L}_2$. Two simple examples contradict this claim:

(1) It is easy to see that a non-injective conservative translation is not enough to preserve intrinsic properties of logics. In fact, any trivial logic[8] $\mathfrak{L}_1$ can be conservatively translated into any other logic $\mathfrak{L}_2$ which merely satisfies $\{b\} \vdash_{\mathfrak{L}_2} b$ for some formula b. The translation mapping $f : \mathfrak{L}_1 \longrightarrow \mathfrak{L}_2$ given by $f(a) = b$ for every formula a of $\mathfrak{L}_1$ is a conservative translation, as it can be easily checked. This means that the triviality of $\mathfrak{L}_1$ is "conservatively" codified in $\mathfrak{L}_2$ by the single inference $\{b\} \vdash_{\mathfrak{L}_2} b$. Additionally, this example shows that trivial theories are not preserved by conservative translations: it suffices to assume in the example above that $\{b\} \nvdash_{\mathfrak{L}_2} b'$ for some formula b' of $\mathfrak{L}_2$. In this case, the trivial logic $\mathfrak{L}_1$ is conservatively translated into the nontrivial theory generated by $\{b\}$. (2) On the other hand, conservative translations do not work well outside the realm of Tarskian logics: if $f : \mathfrak{L}_1 \longrightarrow \mathfrak{L}_2$ is a non-surjective conservative translation such that $\mathfrak{L}_2$ does not satisfy (r3), then the equivalence between formulas of $\mathfrak{L}_1$ is not necessarily preserved by f.

As we shall see later, the failure of conservative translations showed in both examples above can be explained by a meta-level analysis: in both cases, certain metalinguistic existential properties are not preserved by the image of $\mathfrak{L}_1$ under translation f. This can be precisely formalized in the model-theoretic framework to be described now.

The basic idea is to consider a general notion of *attribute* of logics, which formalizes any n-ary relation between formulas and/or sets of formulas. For instance, the consequence relation $\vdash$ of a logic is a relation between sets of formulas and formulas of that logic. On the other hand, any n-ary connective is a (functional) relation between n-uples of formulas and formulas. A mapping between abstract logics (which is called a *transfer*) should preserve attributes; translations between logics are therefore special cases of transfers.

[8] A logic is said to be *trivial* if $\Gamma \vdash a$ holds for any set $\Gamma \cup \{a\}$ of formulas.

From now on, abstract logics are defined as being two-sorted, first-order structures (with one sort for sets of formulas, and the other sort for formulas). On the other hand, transfers are morphisms between such structures, in the sense of Model Theory. Formally:

Definition 3.1 The *basic language of abstract logics* is the first-order two-sorted language $\mathbb{L}$ given by

$$\mathbb{L} = \{\mathbf{form}, \mathbf{Sform}\} \cup \{\varepsilon, \vdash\} \cup \{\Cup, \mathbf{s}\} \cup \{\mathbf{0}\}$$

where $\{\mathbf{form}, \mathbf{Sform}\}$ is the set of basic sorts of $\mathbb{L}$, ε and $\vdash$ are predicate symbols of sort $\mathbf{form} \times \mathbf{Sform}$ and $\mathbf{Sform} \times \mathbf{form}$, respectively; $\Cup : \mathbf{Sform} \times \mathbf{Sform} \longrightarrow \mathbf{Sform}$ and $\mathbf{s} : \mathbf{form} \longrightarrow \mathbf{Sform}$ are function symbols; and $\mathbf{0}$ is a constant of sort $\mathbf{Sform}$. □

As usual, we will write $\tau \; \varepsilon \; \Upsilon$, $\Upsilon \vdash \tau$ and $\Upsilon \Cup \Xi$ instead of $\varepsilon(\tau, \Upsilon)$, $\vdash(\Upsilon, \tau)$ and $\Cup(\Upsilon, \Xi)$, respectively. Variables of sort **form** and **Sform** will be denoted by x, y, z and X, Y, Z (possibly with subscripts), respectively.

The models for the basic language are defined as being the abstract (propositional) logics.

Definition 3.2 An *abstract logic*[9] $\mathfrak{L}$ is a two-sorted structure for the basic language $\mathbb{L}$ of the form

$$\mathfrak{L} = \langle A, P, \varepsilon_{\mathfrak{L}}, \vdash_{\mathfrak{L}}, \Cup_{\mathfrak{L}}, \mathbf{s}_{\mathfrak{L}}, \mathbf{0}_{\mathfrak{L}} \rangle$$

satisfying the following set of axioms in $\mathbb{L}$:

$[Ax1]$ $\quad (\forall X)(\forall Y)(X = Y \Leftrightarrow (\forall x)(x \; \varepsilon \; X \Leftrightarrow x \; \varepsilon \; Y))$;

$[Ax2]$ $\quad (\forall x)(\forall y)(y \; \varepsilon \; \mathbf{s}(x) \Leftrightarrow y = x)$;

$[Ax3]$ $\quad (\forall X)(\forall Y)(\forall x)(x \; \varepsilon \; X \Cup Y \; \Leftrightarrow \; ((x \; \varepsilon \; X) \vee (x \; \varepsilon \; Y)))$;

[9] Please do not confound this notion with that introduced in Definition 2.4 under the same name.

$[Ax4]$ $\quad (\forall x)\neg(x \varepsilon \mathbf{0})$;

$[Ax5]$ $\quad (\forall X)(\exists Y)(\forall x)(x \varepsilon Y \Leftrightarrow \neg(x \varepsilon X))$;

$[Ax6]$ $\quad (\forall X)(\exists Y)(\forall x)(x \varepsilon Y \Leftrightarrow X \vdash x)$.

Let As be the set of axioms $\{[Ax1], \ldots, [Ax6]\}$. If the language $\mathbb{L}'$ is an extension of $\mathbb{L}$, and $\mathfrak{L}$ is a structure for $\mathbb{L}'$, we say that $\mathfrak{L}$ is an *abstract logic for the language* $\mathbb{L}'$ if $\mathfrak{L}$ restricted to $\mathbb{L}$ is an abstract logic. If a logic $\mathfrak{L}$ is a substructure of another logic $\mathfrak{L}'$, we say that $\mathfrak{L}$ is a *sublogic* of $\mathfrak{L}'$. □

The minimal amount of set theory necessary to refer to single formulas, sets of formulas, (finite) unions and the empty set is guaranteed by the satisfaction of the set of axioms As. So, for example, $[Ax1]$ expresses extensionality of sets (in Leibniz's sense), axiom $[Ax2]$ expresses that $\mathbf{s}(x)$ is the singleton set just containing the element x, and $[Ax6]$ establishes that, for each set X, there exists the set of logical consequences of X.

Despite the equality symbol "=" is always interpreted as the identity relation, it is possible to have non-standard models, in which the set P is not necessarily a subset of the powerset $\wp(A)$ of A, or where the interpretation of the singletons $\mathbf{s}(x)$ might not be real singletons, and so on. A model is said to be *standard* if $P \subseteq \wp(A)$, and the symbols ε, $\Cup$, $\mathbf{s}$ and $\mathbf{0}$ are interpreted with its usual set-theoretic sense; in particular, $\mathbf{0}_{\mathfrak{L}}$ is the empty set $\emptyset$. Fortunately, it can be proven that every abstract logic is isomorphic to a standard abstract logic. Because of this, from now on every abstract logic is assumed to be standard. Since in a standard model the mention to the set-theoretic symbols is superfluous, they shall be omitted from now on. Thus, a (standard) abstract logic $\mathfrak{L}$ will be simply denoted by $\mathfrak{L} = \langle A, P, \vdash_{\mathfrak{L}} \rangle$. It is important to observe that P is a Boolean algebra w.r.t. the set-theoretic operations, such that $\emptyset \in P$ and $A \in P$; $\{a\} \in P$ if $a \in A$, and $\{a \ : \ \Gamma \vdash_{\mathfrak{L}} a\} \in P$ if $\Gamma \in P$. An interesting particular case is when $P = \wp(A)$.

Despite the language $\mathbb{L}$ is powerful enough to express meta-properties of logics, it is bounded by the well-known limitations of first-order

logic. In particular, using the *compactness theorem* for first-order logic, it is easy to prove that there is no set $\Psi(X)$ of formulas (just depending on the variable X of sort **Sform**) in any extension $\mathbb{L}'$ of $\mathbb{L}$ with the following property: a (standard) logic $\mathfrak{L}$ satisfies $\Psi(X)$ with parameter Γ iff the set Γ is finite. Thus, meta-properties of logics such as compactness cannot be expressed in any extension of $\mathbb{L}'$.

The notion of *attribute* of a logic and *attribute-preserving mappings* is can be formalized as follows:

Definition 3.3 Let $\mathbb{L}'$ be an extension of $\mathbb{L}$.
(i) An *attribute* of an abstract logic $\mathfrak{L}$ over $\mathbb{L}'$ is just an element of $\mathbb{L}'$.
(ii) Let $\mathfrak{L}_i$ be abstract logics over $\mathbb{L}'$ such that $\mathbf{form}_{\mathfrak{L}_i} = A_i$ and $\mathbf{Sform}_{\mathfrak{L}_i} = P_i$ $(i = 1, 2)$. A *transfer from* $\mathfrak{L}_1$ *into* $\mathfrak{L}_2$ is a morphism $\langle T, T_* \rangle : \mathfrak{L}_1 \longrightarrow \mathfrak{L}_2$ of $\mathbb{L}'$-structures such that $T_*(\Gamma) = T[\Gamma]$ for every $\Gamma \in P_1$. If $\mathbb{L}' = \mathbb{L}$ then a transfer is called a *translation*. □

Since the mapping T_* is redundant in a transfer $\langle T, T_* \rangle : \mathfrak{L}_1 \longrightarrow \mathfrak{L}_2$ (because it is obtained from T), from now on it will be omitted, and we will simply write $T : \mathfrak{L}_1 \longrightarrow \mathfrak{L}_2$.

It is worth noting that, if $\mathbb{L}' = \mathbb{L}$ and $T : \mathfrak{L}_1 \longrightarrow \mathfrak{L}_2$ is a transfer (that is, a translation), then

$$\Gamma \vdash_{\mathfrak{L}_1} a \quad \text{implies that} \quad T[\Gamma] \vdash_{\mathfrak{L}_2} T(a),$$

that is, the usual definition of translation in the sense of Brown & Suszko (1973) is recaptured. Inspired by this, we will say that a transfer $T : \mathfrak{L}_1 \longrightarrow \mathfrak{L}_2$ is *conservative* if, for every $\Gamma \in P_1$ and every $a \in A_1$,

$$\Gamma \vdash_{\mathfrak{L}_1} a \quad \text{if and only if} \quad T[\Gamma] \vdash_{\mathfrak{L}_2} T(a).$$

If $T : \mathfrak{L}_1 \longrightarrow \mathfrak{L}_2$ is a conservative transfer and $P_1 = \wp(A_1)$ then the image $T(\mathfrak{L}_1)$ of $\mathfrak{L}_1$ under T (which is always a substructure of $\mathfrak{L}_2$) is also an abstract logic, that is, it satisfies As (see Definition 3.2).

The approach to abstract logics as model-theoretic structures can take profit from results from Model Theory, obtaining results about the preservation of meta-properties.

Thus, it can be proven that the ultraproduct of a family of abstract logics over a language $\mathbb{L}'$ is a (standard) abstract logic over $\mathbb{L}'$, as a

consequence of Łoś's theorem. More importantly, if a logic $\mathfrak{L}$ satisfies a set $\mathbb{T}$ of Π^0_1-formulas (that is, formulas of the form $(\forall v_1 \cdots \forall v_n)\psi$, where each v_i is a variable of sort **form** or sort **Sform**, and ψ has no quantifiers) then every sublogic $\mathfrak{L}'$ of $\mathfrak{L}$ also satisfies $\mathbb{T}$. Now, consider an abstract logic $\mathfrak{L}_1$ such that $P_1 = \wp(A_1)$. If $T : \mathfrak{L}_1 \longrightarrow \mathfrak{L}_2$ is a transfer then

$$\mathfrak{L}_1 \models \varphi[a_1, \ldots, a_n; \Gamma_1, \ldots, \Gamma_m]$$

implies that

$$\mathfrak{L}_2 \models \varphi[T(a_1), \ldots, T(a_n); T[\Gamma_1], \ldots, T[\Gamma_m]]$$

for every positive formula φ, every $(a_1, \ldots, a_n) \in A_1^n$ and every $(\Gamma_1, \ldots, \Gamma_m) \in P_1^m$.[10] In particular, considering that $T(\mathfrak{L}_1)$ is a logic (whenever T is conservative, c.f. remark above), if $\mathfrak{L}_1$ satisfies a set $\mathbb{T}$ of positive properties then, for every conservative transfer $T : \mathfrak{L}_1 \longrightarrow \mathfrak{L}_2$, the logic $T(\mathfrak{L}_1)$ also satisfies $\mathbb{T}$.

By observing that $\mathbb{L}'$ is a formal metalanguage in which we can express a wide class of meta-properties of logics, the model-theoretic results just mentioned have a deep logical content, showing that this formalism is very appropriate to analyze the meaning of translations between logics.

Abstract logics are very general: no conditions are imposed on the consequence relation $\vdash$. Moreover, no connectives are considered in the set of formulas. In this sense, abstract logics generalize the logics studied by Universal Logic (see Section 2). However, it is easy to restrict our attention to Tarskian logics, by requiring the satisfaction of the following axioms:

[A1] $\quad (\forall Y) Ent(Y, Y)$

[A2] $\quad (\forall Y_1)(\forall Y_2)(\forall x)((Ent(Y_1, Y_2) \wedge (Y_2 \vdash x)) \rightarrow Y_1 \vdash x)$

where $Ent(Y_1, Y_2)$ stands for $(\forall y)((y \,\varepsilon\, Y_2) \rightarrow (Y_1 \vdash y))$. Clearly [A1] and [A2] express, in the formal metalanguage $\mathbb{L}$, properties (r1) and (r3) of Section 2, respectively. Thus, an abstract logic $\mathfrak{L}$ is Tarskian iff satisfies [A1] and [A2].

[10] Recall that a formula φ is *positive* if it has no occurrences of $\rightarrow$ and $\neg$.

On the other hand, if we want to consider logics in which the set of formulas is an algebra generated by a set of connectives (as in the case of "concrete" propositional logics, or as in the case of Brown and Suszko's abstract algebras, recall Definition 2.4), it is enough to add to $\mathbb{L}$ a function symbol $\bar{c} : \mathbf{form}^n \longrightarrow \mathbf{form}$ for each n-ary connective c, obtaining an extension $\mathbb{L}'$ of $\mathbb{L}$. This technique is used, for instance, in abstract algebraic logic, to encode a logic through a (first-order) equational theory (see, for instance, Blok & Pigozzi (1989)). Moreover, given a Hilbert-calculus $\mathcal{H}$, it is straightforward to obtain a set of axioms R in $\mathbb{L}'$ encoding the axioms and inference rules of $\mathcal{H}$. Thus, the least $\mathbb{L}'$-structure satisfying $\mathsf{As} \cup R \cup \{[\text{A1}], [\text{A2}]\}$ is an abstract logic $\mathfrak{L}_{\mathcal{H}}$ encoding the Hilbert calculus $\mathcal{H}$.

Now we are ready to explain the failure of conservative translations in examples (1) and (2) at the beginning of this section. In the case of the first example, consider the formula $\psi(X)$ given by $(\exists y)(X \nvdash y)$. Then $\mathfrak{L}_2 \models \psi(X)[T(\Gamma)]$, but the witness $y \mapsto b'$ does not belong to $T(\mathfrak{L}_1)$. The key is that $T(\mathfrak{L}_1)$ is not an *elementary* substructure of $\mathfrak{L}_2$, that is, T is not an elementary embedding. The second example is justified by the same reason: consider the formula $\phi(x, z)$ given by $(\exists y)((\mathbf{s}(x) \vdash y) \wedge (\mathbf{s}(z) \nvdash y))$. Then it is easy to find logics $\mathfrak{L}_1$ and $\mathfrak{L}_2$, and formulas $a, b \in A_1$ such that $\mathfrak{L}_2 \models \phi(x, z)[T(a), T(b)]$, but the witness $y \mapsto b'$ lies outside the image of T. Again, one explanation for this phenomenon is that $T(\mathfrak{L}_1)$ is not a elementary substructure of $\mathfrak{L}_2$.

But then what is, after all, a "good" notion of translation which preserves "as much as possible" the properties of $\mathfrak{L}_1$ into $\mathfrak{L}_2$? Of course, a "good" translation could be defined as being an isomorphism, but this is quite a strong requirement. Looking at the examples above and the model-theoretical setting presented here, it seems that the notion of *elementary transfer*, that is, a transfer which is an elementary embedding, captures our intuitions at the right level. In fact, the concept of elementary transfer is an intermediate notion between conservative translation and isomorphism, which characterize the transference of "meaning of logics" (given by attributes, recall Definition 3.1) in a good way.

By the very definition, an elementary transfer between logics $\mathfrak{L}_1$ and $\mathfrak{L}_2$ defined in $\mathbb{L}'$ is a transfer $T : \mathfrak{L}_1 \longrightarrow \mathfrak{L}_2$ satisfying the following property: for any formula $\varphi(x_1, \ldots, x_n; X_1, \ldots, X_m)$ in $\mathbb{L}'$ and for any tuples $\vec{a} = (a_1, \ldots, a_n) \in A_1^n$, $\vec{\Gamma} = (\Gamma_1, \ldots, \Gamma_m) \in P_1^m$, the following holds:

$$\mathfrak{L}_1 \models \varphi[\vec{a}; \vec{\Gamma}] \text{ iff } \mathfrak{L}_2 \models \varphi[T(a_1), \ldots, T(a_n); T[\Gamma_1], \ldots, T[\Gamma_m]].$$

As a consequence of this, the logic $\mathfrak{L}_1$ is "faithfully encoded" within $\mathfrak{L}_2$ through T. It should be noted that, as a consequence of the definition, an elementary transfer is an injective mapping, but it is not necessarily a surjective mapping. If T is an elementary transfer then any existential property of the form $(\exists x)\phi$ or $(\exists X)\psi$[11] which is satisfied by $\mathfrak{L}_1$ it must be also satisfied by $\mathfrak{L}_2$ throughout $T(\mathfrak{L}_1)$; that is, there must exist a "witness" in $\mathfrak{L}_2$ of the form $T(a)$ or $T[\Gamma]$, respectively.

Transfers $T : \mathfrak{L}_1 \longrightarrow \mathfrak{L}_2$ such that, for every $\Gamma \in P_1$ and every $a \in A_1$,

$$\Gamma \vdash_{\mathfrak{L}_1} a \quad \text{if and only if} \quad T[\Gamma] \vdash_{\mathfrak{L}_2} T(a)$$

(and, in particular, when $\mathbb{L}' = \mathbb{L}$, conservative translations) are, in fact, a very important measure of *similarity* between $\mathfrak{L}_1$ and $\mathfrak{L}_2$, but this cannot be overestimated. If T is not injective, then example (1) at the beginning of this section shows that a non-injective conservative translation could be useless. Thus, if we restrict ourselves to *injective* conservative transfers, formulas such as $\psi(X)$, given by $(\exists x)\neg(X \vdash x)$, or $\phi(x)$, given by $(\exists X)(X \vdash x)$, contradict the slogan "T encodes $\mathfrak{L}_1$ within $\mathfrak{L}_2$", in case that T is not surjective. Another interesting example (corresponding to example (2) given at the beginning of this section) is the formula $\phi(x, z)$ given by $(\exists y)((\mathbf{s}(x) \vdash y) \wedge (\mathbf{s}(z) \nvdash y))$, mentioned above. In all of these examples, the "witness" for the existential quantifier could be an element of A_2 or P_2, not belonging to $T[A_1]$ or $T_*[P_1]$, respectively (because we are assuming that T is not a surjective mapping). That is, a conservative transfer which fails to be injective or surjective could not translate faithfully $\mathfrak{L}_1$ into $\mathfrak{L}_2$. This means that the only possibility would be to consider bijective

[11] Recall that lower case letters denote variables of sort **form**, whereas capital letters denote variables of sort **Sform**.

conservative transfers. But in the case of the basic language $\mathbb{L}$, this is equivalent to require that T is an isomorphism, quite a strong requirement.

This analysis shows that elementary transfers (in particular, elementary translations), as already noted, constitute an interesting point of equilibrium: they preserve all the relevant properties of $\mathfrak{L}_1$ within the logic $\mathfrak{L}_2$, but the stronger requirement of T being surjective (and, therefore, an isomorphism) is avoided.

Of course we could consider injective conservative transfers preserving some classes of existential formulas. Under this perspective, the more existential formulas are preserved, the "better" the translation. In particular, a basic requirement is to preserve trivialization, in the sense that $T[\Gamma] \vdash_{\mathfrak{L}_2} T(a)$ for all $a \in A_1$ implies that $T[\Gamma] \vdash_{\mathfrak{L}_2} b$ for all $b \in A_2$. In this sense, the "classical" translations of Gödel, Glivenko and Gentzen from classical logic into intuitionistic logic are "good", since they preserve the bottom particle. This a consequence of the following fact: if $\mathfrak{L}_1$ and $\mathfrak{L}_2$ are two (standard) logics over a extension $\mathbb{L}'$ of $\mathbb{L}$ containing a constant $\perp$ such that $\mathfrak{L}_i \models (\forall x)(\mathbf{s}(\perp) \vdash x)$ $(i = 1, 2)$, and if $\mathfrak{L}_2$ satisfies

$$(\forall Y_1)(\forall Y_2)(\forall x)(\forall y)[((Y_1 \vdash x) \ \wedge \ (Y_2 \Cup \mathbf{s}(x) \vdash y)) \rightarrow (Y_1 \Cup Y_2 \vdash y)]$$

(the *cut property*) then, for any transfer $T : \mathfrak{L}_1 \longrightarrow \mathfrak{L}_2$ and any $\Gamma \in P_1$, the following holds: $T[\Gamma] \vdash_{\mathfrak{L}_2} T(a)$ for all $a \in A_1$ implies that $T[\Gamma] \vdash_{\mathfrak{L}_2} b$ for all $b \in A_2$.[12]

By observing this kind of examples, it would be interesting, as suggested above, to think about a wide range of increasingly "better" notions of translations, depending on the amount of existential formulas that are preserved.

We hope that the discussion about conservativeness of attributes of logics through translations, under the model-theoretic perspective presented in this section, can throw some light to the question about the meaning of translations between logics. In the next section we

[12]Observe that, since $\perp$ is a constant, $T(\perp_{\mathfrak{L}_1}) = \perp_{\mathfrak{L}_2}$. From this the result follows easily.

will analyze the importance of preserving meta-properties of a special kind by translations, in the realm of combination of logics. This analysis provides additional evidence in favor of the thesis proposed here: (most of) meta-properties of a logic should be preserved by a logic translation.

4. META-TRANSLATIONS FOR COMBINING LOGICS

In recent years the subject of combination of logic systems has grown up considerably, and a number of news methods have been developed, whereas techniques already established were improved.

In general, if $\mathcal{L}_1 \oplus \mathcal{L}_2$ is the logic system obtained by combining logics $\mathcal{L}_1$ and $\mathcal{L}_2$ (using some method), then it is expected that $\mathcal{L}_1 \oplus \mathcal{L}_2$ should be the "least" logic system (under some measure) that "extends" both $\mathcal{L}_1$ and $\mathcal{L}_2$ (in a certain well-defined sense). The concrete meaning of the expressions between quotation marks depends on each combination method.

As a consequence of this, it seems clear that, in general, a method for combining logics presupposes or requires at least two tasks:

(i) to represent logics systems in a general way; and (ii) to have a good theory of translations between logics, able to formalize the notion of extension. This is why the topic of combining logics is relevant to our discussion here.

Translations between logics, in the sense of Brown and Suskzo, are usually employed for any method for combining logics. In the case of decomposing logics into simpler logics, translations are fundamental for the technique known as *possible-translation semantics*, introduced in Carnielli (1990). The basic idea of this method is to analyze a given logic $\mathcal{L}$, which is not well-known or it is hard to deal with, by means of translations $f_i : \mathcal{L} \longrightarrow \mathcal{L}_i$ $(i \in I)$, where each $\mathcal{L}_i$ is a logic generally simpler than $\mathcal{L}$. The basic property of each translation f_i is the usual: the mapping f_i must preserve the consequence relation. At the linguistic level, these translations are frequently defined as in Definition 2.6, but sometimes they are a bit more complicated. See, for instance, Marcos (forthcoming) for some examples of possible-translations semantics.

On the other hand, there are methods for composing logics, that is, for obtaining a logic $\mathcal{L}_1 \oplus \mathcal{L}_2$ as a result of combining $\mathcal{L}_1$ and $\mathcal{L}_2$, as it was mentioned at the beginning of this section. *Fibring*, introduced in Gabbay (1996), is one of the most developed techniques for composing logics. From the seminal paper Sernadas, Sernadas & Caleiro (1999), this method can be expressed in terms of Category Theory as a coproduct of the given logics, computed in the category in which the logic systems are being represented.[13] Independent of the category chosen for representing the logic systems (consequence relations in the sense of Definition 2.4, Hilbert calculi, logics presented semantically etc.), the notion of morphism between logics systems is required to satisfy the same basic property: all of them must preserve the logical inferences.

At first sight, it seems reasonable to adopt this point of view: if we want to define the least (conservative) extension $\mathcal{L}_1 \oplus \mathcal{L}_2$ of the logics $\mathcal{L}_1$ and $\mathcal{L}_2$, then the inclusion maps $f_i : \mathcal{L}_i \longrightarrow \mathcal{L}_1 \oplus \mathcal{L}_2$ should be (conservative) translations. This is rather obvious. If we recall the model-theoretic analysis of meta-properties preserved by logic morphisms made in Section 3, this is equivalent to say that the morphisms to be considered preserve (positive) meta-properties of the form $\Gamma \vdash \varphi$. Of course the preservation of *negative* meta-properties of the form $\Gamma \nvdash \varphi$ is, in general, inappropriate for combining logics: it is absolutely plausible to have a situation in which $\Gamma \nvdash_{\mathcal{L}_1} \varphi$ in the logic $\mathcal{L}_1$, whereas $\Gamma \vdash_{\mathcal{L}_1 \oplus \mathcal{L}_2} \varphi$ in the richer combined logic $\mathcal{L}_1 \oplus \mathcal{L}_2$ (note that, in this case, $\mathcal{L}_1 \oplus \mathcal{L}_2$ is a non-conservative extension of $\mathcal{L}_1$).

[13] This is the case of the so-called *unconstrained fibring*. There is another kind of fibring called *constrained fibring*, in which some connectives of the given logics are allowed to be shared. This operation can also be described in terms of Category Theory but, by simplicity, and since this goes out of the scope of this paper, we will not give the details here.

In order to keep things simple just consider two categories of logic systems:[14] **Hil** and **Con**, formed by logic presented by Hilbert calculi and (Tarskian, structural and compact) consequence relations, respectively. Both categories are based on a fixed category **Sign** of signatures, and so a logic is a pair $\langle \Sigma, R\rangle$ (in the case of **Hil**) or a pair $\langle \Sigma, \vdash\rangle$ (in the case of **Con**) such that Σ is a signature (recall the discussion just before Definition 2.6), that is, an object of **Sign**; R is a set of inference rules over the propositional language $L(\Sigma)$, and $\vdash$ is a Tarskian, structural and compact consequence relation defined over $L(\Sigma)$.[15] At the linguistic level, the morphisms in **Sign** are generally defined as being homomorphisms between the algebras of formulas, as in Brown and Suszko's approach (recall Definition 2.5), or can be a bit more elaborated, as in Wójcicki's approach (recall Definition 2.6). As mentioned above, a morphism between logics in **Hil** or in **Con** is a morphism in **Sign** which preserves the logic inferences, that is, a translation in the sense of Brown and Suszko. Thus, an injective morphism $i : \mathcal{L}_1 \longrightarrow \mathcal{L}_2$ in **Hil** or in **Con** is an injective morphism $i : \Sigma_1 \longrightarrow \Sigma_2$ in **Sign** between the respective signatures such that: $\Gamma \vdash_{\mathcal{L}_1} \varphi$ implies $i[\Gamma] \vdash_{\mathcal{L}_2} i(\varphi)$. In other words, the logic $\mathcal{L}_2$ is a (not necessarily conservative) extension of $\mathcal{L}_1$, through the injective syntactic translation i. As mentioned above, the fibring of two logics systems (in **Hil** or in **Con**) is the coproduct of the systems (computed in **Hil** or in **Con**, respectively) and the signature of the obtained system is the coproduct (in **Sign**) of the respective signatures. As a matter of fact, the coproduct in **Con** is given by the supremum of (the embedding of)

[14] In the sequel, me will use, again, some basic terminology from Category Theory. The reader not familiar with these terms just can ignore them, because they are not strictly necessary for understanding the main ideas underlying the arguments.

[15] Note the difference between this approach and the notions of institutions and entailment systems (recall Definitions 2.7 and 2.8): in the latter, each individual logic owns a category of signatures ("notations"), whereas in the present approach there is a fixed category of signatures, and each logic system is *presented* in a single signature (object of **Sign**).

the given consequence relations, computed in the (complete) lattice of consequence relations over the coproduct signature.

In Coniglio (forthcoming) it was observed the following phenomenon: suppose that we try naturally to recover classical logic, defined over the language consisting of negation and disjunction, from its basic logical component (that its, the rules for negation, on the one side, and the rules for disjunction, on the other), by combining both components using fibring. Surprisingly, the result in both **Hil** and **Con** is a logic weaker than classical logic, in which the formula $(\varphi \vee \neg\varphi)$ is not a theorem. In the case of **Hil**, this result is independent of the (sound and complete) specific axiomatization chosen for each connective.

In the same article were given another examples of this phenomenon. For instance, it is impossible to obtain, by fibring in **Hil** or in **Con**, classical logic by combining the logic of classical negation with the logic of classical implication, against the expectations. In this case, the formula $\varphi \Rightarrow (\neg\varphi \Rightarrow \psi)$ does not hold in the resulting logic system. Moreover, the deduction meta-theorem is no longer valid in the resulting logic: the formula ψ is always derivable from $\{\varphi, \neg\varphi\}$, despite there are instances of $\varphi \Rightarrow (\neg\varphi \Rightarrow \psi)$ which are not derivable, as it was just mentioned. This means that certain (positive) meta-properties of the given logics are missing through the process of fibring.

All the examples above show that the logics obtained by fibring are, in some sense, too *weak*: some expected formulas cannot be derived. This problem arises mainly because the only attribute of the logics (recall Definition 3.3) that is preserved by fibring is the consequence relation: if $\Gamma \vdash_{\mathcal{L}_1} \varphi$ then $\Gamma \vdash_{\mathcal{L}} \varphi$ holds in the logic $\mathcal{L} = \mathcal{L}_1 \oplus \mathcal{L}_2$ obtained by fibring. As discussed in Section 3, a (meta)formula of the form $\Gamma \vdash_{\mathcal{L}_1} \varphi$ is a meta-property of $\mathcal{L}_1$, a basic one indeed. But, for instance, a meta-property such as the deduction meta-theorem

$$\Gamma, \varphi \vdash_{\mathcal{L}_1} \psi \quad \texttt{iff} \quad \Gamma \vdash_{\mathcal{L}_1} \varphi \Rightarrow \psi$$

is a more complex meta-property of $\mathcal{L}_1$, that would be also preserved by fibring. However, as it was showed, frequently this is not the case.

In the case of the combination of negation with disjunction, the meta-property

$$\frac{\Gamma, \varphi \vdash \psi \quad \Delta, \neg\varphi \vdash \psi}{\Gamma, \Delta \vdash \psi}$$

of the consequence relation $\vdash$ associated to the logic of classical negation was not preserved by the fibring based on usual translations. The inclusion (mono)morphism $i : \mathcal{L}_1 \longrightarrow \mathcal{L}_1 \oplus \mathcal{L}_2$ above mentioned just preserves the consequence relation, that is, basic facts of the form "$\Gamma \vdash \varphi$", and not more complex statements (recall the formalized metalanguage for logics described in Section 3). In Coniglio (forthcoming) were introduced categories of deduction systems (generalizing sequent calculi) in which the morphism $h : \mathcal{L} \to \mathcal{L}'$ preserve meta-properties of the logics of the form

$$\texttt{If } \Gamma_1 \vdash_{\mathcal{L}} \varphi_1 \texttt{ and } \ldots \texttt{ and } \Gamma_n \vdash_{\mathcal{L}} \varphi_n \texttt{ then } \Gamma \vdash_{\mathcal{L}} \varphi.$$

In other words: from a meta-property of $\mathcal{L}$ of the form

$$\texttt{If } \Gamma_1 \vdash_{\mathcal{L}} \varphi_1 \texttt{ and } \ldots \texttt{ and } \Gamma_n \vdash_{\mathcal{L}} \varphi_n \texttt{ then } \Gamma \vdash_{\mathcal{L}} \varphi$$

the following meta-property of $\mathcal{L}'$ must be deduced:

$$\texttt{If } h[\Gamma_1] \vdash_{\mathcal{L}'} h(\varphi_1) \texttt{ and } \ldots \texttt{ and } h[\Gamma_n] \vdash_{\mathcal{L}'} h(\varphi_n)$$
$$\texttt{then } h[\Gamma] \vdash_{\mathcal{L}'} h(\varphi).$$

As a consequence of this, when a logic system is embedded in a larger one by fibring then any meta-property (of the form above) is preserved by the canonical injection, by the very definition of morphism. The resulting operation is accordingly called *meta-fibring*, and the morphisms preserving this kind of meta-properties are called *meta-translations*.

Observe that a meta-translation is weaker than an elementary transfer, that is, an elementary embedding with respect to the formalized metalanguages described in Section 3. On the other hand, it is stronger than the usual notion of translation (or transfer): usually,

a first-order morphism just preserves positive formulas, and the meta-properties preserved by a meta-translation are not positives, because they contain a (meta)implication. The useful applications of meta-translations to fibring theory show that meta-translations constitute an interesting intermediate notion between usual translations and elementary transfers.

Another interesting example of the different results that can be obtained by fibring using meta-translations in the place of translations can be found in Béziau & Coniglio (2005). In this paper it was shown that, in **Con**, the fibring of the logic $\mathcal{L}_1$ of classical conjunction with the logic $\mathcal{L}_2$ of classical disjunction produces a logic (over the signature $\{\wedge, \vee\}$) which is not distributive, that is, where the meta-property

$$\varphi \wedge (\psi_1 \vee \psi_2) \vdash (\varphi \wedge \psi_1) \vee (\varphi \wedge \psi_2)$$

is not valid. The proof of this fact is simple: the consequence relation of the logic obtained by fibring in **Con** is the infimum (in the complete lattice of the Tarskian, compact and structural consequence relations over signature $\{\wedge, \vee\}$) of the set of all the consequence relations *containing* the valid inferences concerning classical conjunction and disjunction.[16] The logic of lattices is one of them, however this logic is not distributive (because there exists non-distributive lattices), so the logic obtained by fibring cannot be distributive.

The explanation is the same that in the other examples of this section: the meta-property

$$\texttt{If } \Gamma, \psi_1 \vdash_{\mathcal{L}_2} \varphi \texttt{ and } \Delta, \psi_2 \vdash_{\mathcal{L}_2} \varphi \texttt{ then } \Gamma, \Delta, \psi_1 \vee \psi_2 \vdash_{\mathcal{L}_2} \varphi$$

of the logic $\mathcal{L}_2$ of classic disjunction does not necessarily hold in a given logic extending $\mathcal{L}_2$: in order to obtain the fibring where considered all the logics containing all the inferences (meta-properties of the form $\Gamma \vdash_{\mathcal{L}_2} \varphi$) of $\mathcal{L}_2$, and not more sophisticated properties of $\mathcal{L}_2$. However, if the usual notion of translation between logics is substituted by meta-translations, the distributivity law is recovered.

[16] This is the way the supremum of two logics is computed in this lattice.

This shows that the same operation (coproduct, in this case) produces very different results when the underlying notion of translation between logics is changed. It should be noticed that the same result about the non-derivability of the distributive law by fibring using translations can be obtained in the category **Hil** (see Coniglio (forthcoming)). That is, there are no (sound and complete) axiomatizations for the logics of classical conjunction and disjunction which allow to recover distributivity when using standard translations for combining both systems.

Of course, it is possible to argue against the obtainment of distributivity by combining conjunction with disjunction: after all, this is a new law of the mixed language and this could be seen as an unexpected interaction (see Béziau (2004)). However, there are many contexts (for instance, the question of recovering a logic from its basic components) that justify the requirement of obtaining new interactions between the connectives.

In general, some natural interactions between the rules or axioms defining the connectives of the given logics are to be expected in the combined logic. Any process of *combination* presupposes some kind of *interaction* between the factors, and not the mere adjunction of them (chemical processes are a good analogy here). All the examples given above are characterized by the loss of some expected interactions. In the case of the combination of negation with disjunction, the interaction law $(\varphi \vee \neg\varphi)$ is not obtained, whereas $\varphi \Rightarrow (\neg\varphi \Rightarrow \psi)$ is the interaction law that is not valid in the combination of negation with implication.

To conclude, the examples given in this section from the area of combination of logics support the idea that a stronger notion of translation between logics should be based on the preservation of meta-properties stronger than the usual ones.

5. CONCLUDING REMARKS

This article reviewed several approaches to the question of the representation of logic systems and their mappings, that is, the logic translations.

From this brief analysis, it was supported the thesis that a notion of translation between logics stronger than the usual one should be based on the preservation of meta-properties. In Section 3 we saw that a model-theoretic approach seems appropriate for formalizing the meta-language in which the meta-properties of logic systems can be expressed. Moreover, the notion of *elementary transfer* was found to be an interesting notion of mapping which allows to "transfer" (first-order) meta-properties of a logic into another, without requiring to be an isomorphism.

Finally, in Section 4 we saw that the preservation of certain class of meta-properties characterize a notion of translation (called *meta-translations*) which is useful when applied to mechanisms for combining logics.

Of course it is possible to disagree with some of the conclusions we arrive. However, it is interesting to observe that the introduction of a very general formal concept (in this case, translations between logics) is connected with the applications we have in mind. The first definition of translation between logics given by Glivenko today sounds very weak for most applications; nevertheless, for the applications that it was created (the proof of consistency of a logic relative to the consistency of another logic) this definition was enough. Analogously, if we want to recover a logic from its components, the usual notion of translation between logics could not be enough, and the use of meta-translations would be more appropriate. We hope that this discussion can help to understand what is the meaning of a translation between logics in a deeper sense.

REFERENCES

BELL, J.L. *Toposes and local set theories: an introduction.* Oxford: Clarendon Press, 1988.

BÉZIAU, J.-Y. *Recherches sur la Logique Universelle (Excessivité, Négation, Séquents).* Ph.D. Thesis, Paris 7, 1994.

BÉZIAU, J.-Y. "Les Axiomes de Tarski". Forthcoming. Available at http://www.unine.ch/unilog/jyb/Les%20axiomes%20de%20Tarski.pdf

BÉZIAU, J.-Y. "A paradox in the combination of logics". In: W.A. Carnielli, F.M. Dionísio, and P. Mateus, (eds.) (2004), pp. 87-92.

BÉZIAU, J.-Y., CONIGLIO, M.E. "Combining Conjunction with Disjunction". In: B. Prasad (ed.) (2005).

BLOK, W., PIGOZZI, D. *Algebraizable Logics*, volume 77 (396) of *Memoirs of the American Mathematical Society*. Providence, RI: AMS, 1989.

BORZYSZKOWSKI, A.M., SOKOLOWSKI, S. *Proc. of Mathematical Foundation of Computer Science '93*. Number 711 in *Lecture Notes in Computer Science*. Berlin: Springer Verlag, 1993.

BROWN, D.J., SUSZKO, R. "Abstract Logics". *Dissertationes Mathematicae* 102, pp. 9-41, 1973.

CAICEDO, X., MONTENEGRO, C. *Models, Algebras and Proofs – Selected papers of the X Latin American symposium on mathematical logic held in Bogotá*. Volume 203 in *Lectures Notes in Pure and Applied Mathematics*. New York: Marcel Dekker Inc., 1999.

CARNIELLI, W.A. "Many-valued logics and plausible reasoning". In *Proceedings of the XX International Congress on Many-Valued Logics*, pp. 328–335. IEEE Computer Society, 1990.

CARNIELLI, W.A., DIONÍSIO, F.M., MATEUS, P. *Proceedings of Comblog'04*. Lisbon: IST, 2004.

CERIOLI, M., MESEGUER, J. "May I Borrow Your Logic?". In: A.M. Borzyszkowski and S. Sokolowski (eds.) (1993), pp. 342-351.

CONIGLIO, M.E., CARNIELLI, W.A. "Transfers between logics and their applications". *Studia Logica*, 72(3), pp. 367-400, 2002.

CONIGLIO, M.E. "The meta-fibring environment: preservation of meta-properties by fibring". Forthcoming. A preliminary version was published at *CLE e-Prints*, 5(4), 2005. Available at http://www.cle.unicamp.br/e-prints/vol_5,n_4,2005.htm

DA SILVA, J.J., D'OTTAVIANO, I.M.L., SETTE, A.M. "Translations between Logics". In: X. Caicedo and C. Montenegro (eds.) (1999), pp. 435-448.

D'OTTAVIANO, I.M.L., FEITOSA, H.A. "Many Valued Logics and Translations". *The Journal Of Applied Non Classical Logics*, v. 9, n. 1, pp. 121-140, 1999.

D'OTTAVIANO, I.M.L., FEITOSA, H.A. "Conservative Translations". *Annals of Pure and Applied Logic*, v. 108, pp. 205-227, 2001.

EBBINGHAUS, H.D., FERNANDEZ-PRIDA, J., GARRIDO, M. *Logic Colloquium '87: Proceedings of the Colloquium Held in Granada, Spain July 20-25, 1987*. 1989.

FEITOSA, H.A. *Traduções Conservativas* (*Conservative Translations*, in Portuguese). Ph.D. Thesis, State University of Campinas, Brazil, 1997.

GABBAY, D. "Fibred semantics and the weaving of logics: Part 1". *The Journal of Symbolic Logic*, 61(4), pp. 1057-1120, 1996.

GOGEN, J.A., BURSTALL, R.M. "Introducing Institutions". In: *Logics of Programs (Carnegie-Mellon University, June 1983)*, pp. 221-256. Volume 164 in *Lecture Notes in Computer Science*. Springer, 1984.

GOGEN, J.A., BURSTALL, R.M. "Institutions: Abstract model theory for specification and programming". *Journal of the ACM*, 39, n. 1, pp. 95-146, 1992.

HERTZ, P. "Über Axiomensysteme für beliebige Satzsysteme". *Matematische Annalen*, 101, pp. 457-514, 1929.

ŁOŚ, L., SUSZKO, R. "Remarks on sentential logics". *Proc. Kon. Akad. van Weten.*, ser. A., 61 and IM 20, pp. 177–183, 1958.

MARCOS, J. "Possible–translations semantics for some weak classically-based paraconsistent logics". Forthcoming in *Journal of Applied Non-Classical Logics*.

MESEGUER, J. "General Logics". In: H.D. Ebbinghaus, J. Fernandez-Prida, M. Garrido and D. Lascar (eds.) (1989), pp. 275-329.

PRASAD, B. *Proceedings of the 2nd Indian International Conference on Artificial Intelligence, IICAI 2005*. Pune: IICAI, 2005.

SERNADAS, A., SERNADAS, C., CALEIRO, C. "Fibring of logics as a categorical construction". *Journal of Logic and Computation*, 9(2), pp. 149-179, 1999.

TARSKI, A. *Logics, Semantics, Metamathematics - Papers from 1923 to 1938*. Oxford: Oxford University Press, 1956.

WÓJCICKI, R. *Theory of Logical Calculi*. Dordrecht: Kluwer Academic Publishers, Synthese Library, 1988.