Consequence Mining: A New Approach to Logical Constants*

Denis Bonnay
Université Paris Ouest Nanterre

Dag Westerståhl
University of Gothenburg

May 9, 2010

1 Introduction

Why is the concept of a logical constant important? Surely the strongest reason is that it appears to be involved in any attempt to identify a precise notion of *logical consequence*, and related notions like logical truth and analyticity. Whether you approach logical consequence semantically, in terms of preservation of truth, or syntactically, in terms of derivability in a formal system, you first need to make a selection of logical constants, which will play a special role in your definition of consequence, in contrast with other symbols which can be treated schematically.

1.1 The problem of logicality

The recognition that different words play different roles for the consequence relation is an old one. An early expression of it was the medieval distinction between *categorematic* and *syncategorematic* terms, the dominant idea being that the latter had no independent meaning (did not correspond to anything 'definite' in the mind). This particular idea is hard to defend in present day semantics, where familiar logical constants like *every* and *or* do receive seemingly independent interpretations of their own. And similarly straightforward modern dichotomies, like the one in linguistics between *functional* and *descriptive* terms, don't seem to give the right result either.[1]

Of course, it is not necessary to have a characterization of logicality before doing logic. Most presentations of formal systems, and their associated notion of consequence, simply start by choosing some familiar logical words, or inventing some new ones, without giving other reasons than, possibly, various desirable properties of the resulting system. It is only when the target of interest is

*Various thanks TBW.

[1] For example, in type-theoretic semantics, all types except the basic ones are functional. And even in higher types there seem to be plenty of non-logical words.

the concept of logical consequence itself, that the need for some independent grounds for that choice makes itself felt.

When Tarski gave the first version of the model-theoretic definition of logical consequence, in Tarski (1936), he plainly recognized that you need such grounds — and that he had none. In fact, this recognition itself is a non-trivial intellectual achievement. The same thing was recognized by Bolzano — whose investigations of logical consequence in the *Wissenschaftslehre* were in some respects more wide-ranging and penetrating than Tarski's (as we will see shortly) — one hundred years earlier.[2] But apart from those two cases, recognition of the problem of logical constants, let alone analyses of logicality — of precisely what it is that distinguishes words like *every* from words like *man*, for example — have been rare until fairly recently.

One reason for this state of affairs was presumably that for a long time, the 'candidate' logical constants seemed few and familiar. But with the advent of generalized quantifier theory in the 1950's and 60's, there suddenly appeared a huge number of expressions about which it makes sense to ask whether they are logical or not. Likewise, the present day explosion of logical systems for various purposes (think of intensional logics, dynamic logics, game-theoretic logics, linear logic, etc.) has also significantly increased the range of putative logical expressions. In any case, there are now some rather well developed attempts on the market to explicate the notion of logicality.

Perhaps the most successful of these to date are characterizations in terms of *invariance* under (a group of) transformations. Starting with Tarski (1986), the plan is to characterize logical notions as the ones invariant under the most general kind of transformation, thereby capturing the idea that logic is the most general of sciences. This field has been rather active recently.[3]

1.2 Constants from consequence: Bolzano reversed

In the present paper, however, we shall approach the notion of a logical constant from a different, and as far as we know, unexplored direction.[4] In a nutshell, the idea is to combine Bolzano's insight that any choice of constants determines a semantic consequence relation with a method for going in the opposite direction:

[2]It is only in one short paragraph in §148 of Bolzano (1837) that he mentions the problem and connects it to defining what he calls logical analyticity. According to Bar-Hillel (1950), this was "surely one of the most important and ingenious single logical achievements of all times." (p. 101) Bar-Hillel notes that this notion of logical truth or analyticity is not used in the rest of the book, and hypothesizes that it was a late insight, inserted just before publication. Be that as it may, an equally important achievement in *Wissenschaftslehre*, which pervades the book, is the analysis of *consequence*, and the realization that it is relative to a selection of constants; see Section 3 below.

[3]See, for example, Bonnay (2008) and Feferman (2010), and the references cited therein, for recent surveys and contributions in this area. For a non-technical overview of various approaches to logical constants, see MacFarlane (2009).

[4]The only exception is Carnap. He did address the question in Carnap (1937), suggesting that the set of logical constants be the maximal set of expressions such that every sentence built out of these expressions is either valid or invalid. Our proposal is essentially different, and closer in spirit to standard accounts of validity.

to *extract*, from any given consequence relation, its constants. The method might thus be called *consequence mining*. We see this as a complement rather than an alternative to the invariance approach. The two approaches focus as it were on different aspects of logical constants. Very roughly, invariance targets *logicality*: the generality and formality of logic. Our approach here targets *constanthood*, in the sense of what must be held constant for an inference to be valid. This goes beyond pure logic, since it applies to any given notion of consequence, of what 'follows from' what, in formal or natural languages.

How can one 'extract the constants' if nothing but a language and a consequence relation is given? The idea is embarrassingly simple. In a given inference which is valid according to that relation, it is often perfectly clear which words or symbols are constant and which are not: Just replace them by others of the same category and see what happens. If validity (according to the given consequence relation) is always preserved under such replacement, we are not dealing with a constant. But if validity can be destroyed in this way, we are. In other words, the constants of that inference are those words or symbols that are *essential* to its validity, in the sense indicated. Here is an example.

(1) a. Most French movies encourage introspection
b. All movies which encourage introspection are commercial failures
c. *Hence:* Most French movies are commercial failures

(1) is presumably valid according to most natural notions of consequence: the conclusion *follows from* the premises. Now replace words like *French*, *movies*, etc. with others of the same category (in some suitable sense):

(2) a. Most red sports cars are convertibles
b. All cars which are convertibles are unsuitable for cold climates
c. *Hence:* Most red sports cars are unsuitable for cold climates

Nothing happens. The inference is still valid. Indeed, in a sense it is the *same* inference. But try instead to replace words like *most* or *all*:

(3) a. No French movies encourage introspection
b. All movies which encourage introspection are commercial failures
c. *Hence:* No French movies are commercial failures

This is not only invalid: whether it is valid or not seems to have nothing to do with the validity of (1)! Hence, *most* is a constant in (1), *French* is not.

These observations are so obvious that they are easy to overlook. But that doesn't make them trivial. The method of replacing words by others of the same category, and seeing what happens, was precisely Bolzano's approach to consequence. In his case, the criterion was that *truth* should be preserved (for consequence to hold). In the reverse direction, we ask instead if *validity* is preserved.

So far, however, we have only given a way to detect constanthood relative to a particular valid inference. The next task is to extract the constants from the consequence relation itself. The present approach to constanthood was first proposed in Peters and Westerståhl (2006), Ch. 9. The suggestion there was that a symbol u is constant if *every* valid inferences in which it occurs essentially can be destroyed by some replacement. The problem with this idea is that the qualification "in which it occurs essentially" is crucial, and must be explained independently, as the following examples from first-order logic illustrate:

(4) a. $Pa \models Pa \vee \exists x Rx$

b. $\exists x Px, \forall x(Px \leftrightarrow Rx) \models \exists x Rx$

In both of these inferences, the quantifier $\exists$ can be replaced by any type $\langle 1 \rangle$ quantifier Q, i.e. with any symbol of the same category, without destroying validity. Cases like (4a) are easy to set aside: the occurrence of $\exists$ there is *spurious*, in that (4a) is an instance of a (in a precise sense) more general inference (i.e. $\varphi \models \varphi \vee \psi$) in which $\exists$ does not occur. But (4b) is more tricky: it expresses a principle of *extensionality*, and it is not clear how all such principles can be set aside on syntactic or other grounds. In particular, the suggested notion of constanthood makes it a non-trivial task to verify that the usual logical symbols in familiar consequence relations, such as first-order consequence, are indeed constants according to the criterion.

Here we shall use a weaker and, as it turns out, more satisfactory criterion: It suffices, for a word or symbol to be constant relative to the given consequence relation, that there be *at least one* valid inference which can be destroyed by a replacement.

To repeat, the main innovation is extraction (mining): starting from a *given* consequence relation, we identify its constants. There is an intuitive appeal to this idea, we feel, which is lacking in abstract discussions of logicality. Often, speakers have rather clear judgments about 'what follows from what', at least much clearer than about the theoretical issue of which words are logical. Our proposal allows us to use only judgments about validity and non-validity of inferences to identify the constants.

1.3 Two criteria of adequacy

Although our general idea is fairly clear, there are choices to be made about the details of its implementation. Moreover, there doesn't seem to exist any similar investigations, or any accepted requirements that such an investigation should satisfy. In order to have some guarantee that we are on the right track, we shall be guided by two criteria. The first is straightforward: The method of extraction should yield the expected logical constants when applied to familiar and well-defined notions of logical consequence.

The second criterion is more vague, but still useful as a guide. Bolzano's definition of consequence is unique in its generality. He realized that *any* set X of words could in principle be selected as constants, and logical consequence with respect to X be defined in a uniform way: preservation of truth however the words *outside* X are replaced. Some choices of X may yield unintuitive results, others may give rise to interesting consequence relations, but the definition is completely general. In effect, he defines a function from arbitrary sets of words to consequence relations. Now, extraction of constants is a function in the opposite direction, from consequence relations to sets of words. But the method of abstraction is based on the same fundamental idea: to study what happens to truth and validity, respectively, under appropriate replacements of words. Therefore, we would hope for the extraction function to be an *inverse* of Bolzano's function, in some suitable sense.

This second guiding principle turns out to be quite fruitful. We will see what kind of inverse relationship can be expected, and get to know in some detail under what conditions on the language and on the notion of consequence it holds. In fact, most of the technical work to follow was driven by the goal of understanding the relationship between the two functions.

1.4 Plan

Our aim is to draw an abstract mathematical picture of the situation just indicated for a given language, with the two functions, one forming consequence relations from constants, and the other extracting constants from consequence relations. We make as few assumptions as possible about the language. But we start from Bolzano's substitutional perspective rather than the model-theoretic one: there is a given *interpreted* language, in which meaningful expressions can be *replaced*, rather than *reinterpreted*. The reasons for this are twofold, but they are not historical.

A first reason is that while there are well-known facts about how lack of expressivity in the language may affect the consequence relations (for example, by inability to express an existing counter-example to an inference), there are analogous effects on the extraction function which are not well-known. It will be instructive to study such effects on both sides in some detail, and to see precisely what is required to avoid them. So our procedure will be to add requirements successively, eventually arriving at a situation where all unwanted effects have been removed. In that situation, the Bolzano style account will be equivalent to the Tarskian account for interpreted languages[5], so the results obtained will transfer to it, but we will have obtained several insights along the way.

The second and main reason, however, for not starting in a model-theoretic setting is that it is not so clear what an arbitrary consequence relation in an *uninterpreted* language is, or if such a notion even makes sense. In standard model-theoretic accounts, languages are *partially* interpreted: the logical symbols have fixed interpretations, the others not. But for our purposes, we need to start without any such distinction: all symbols must be treated equally, and the task is precisely to use properties of the consequence relation to single out its constants. On the other hand, in an interpreted language the concept of an arbitrary consequence relation is straightforward and unproblematic. In fact,

[5]The model-theoretic definition of logical consequence is given for (partly) uninterpreted languages. However, the semantic account is of course not tied to the use of uninterpreted languages. Tarski's original definition of logical consequence in Tarski (1936) is given for an interpreted language.

when we eventually come to the Tarskian semantic account, we will see that languages must still be treated as interpreted, although there is a mechanism for arbitrarily *changing* the interpretation.

Thus, we begin, after preliminary definitions of the notions of *language*, *replacement*, and *consequence relation* to be used (Section 2), with a precise account of Bolzano's function from sets of symbols to consequence relations, and some of its properties (Sections 3 and 4). We then introduce in Section 5 the extraction function in the other direction, and see how it behaves on certain examples, compared to how we would like it to behave, in view of our two adequacy criteria. It turns out to be the second criterion that requires additional assumptions, and the rest of the paper investigates the effects of these. The form of inverse relationship that we find is that of a *Galois connection* (Sections 6 and 7). It holds under certain requirements, in particular in our most general version of the substitutional framework, which essentially is equivalent to a model-theoretic one (Section 8). We end with some concluding remarks and directions for further study.

2 Preliminaries

2.1 Languages

In the Bolzano setting with an interpreted language, we shall take every sentence to be either *true* or *false*. We need very few assumptions about what sentences look like or how they are structured. Most of what we need is captured in the following definition.

Definition 1
An (interpreted) *language* is a triple

$$L = \langle \mathit{Symb}_L, \mathit{Sent}_L, \mathit{Tr}_L \rangle,$$

where

(i) Symb_L is a countable set of atomic *symbols*;
(ii) Sent_L is a set of *sentences*, which are finite strings of signs, some of which belong to Symb_L;[6]
(iii) Tr_L, the set of *true* sentences, is a subset of Sent_L.

As long as we only consider replacement of symbols by other symbols, we may disregard finer aspects of the structure of sentences, such as tree structure. But since we cannot realistically expect a symbol to be meaningfully replaceable by *any* other symbol, we shall presuppose a partition of symbols into categories. More precisely, let a set *Cat* of *categories* be given. We assume that for each language L,

$$\mathit{Cat}_L = \{C_L : C \in \mathit{Cat}\}$$

[6]Think of the other signs as grammatical morphemes, parentheses, commas, variables, etc. We assume there are at most countably many of those too, and then Sent_L is also countable.

is a *partition* of $Symb_L$. We say that $u \in Symb_L$ *is of category* C if $u \in C_L$, and we allow that C_L is empty when L has no symbols of category C. Note that symbols in distinct languages can be of the same category.

Whenever convenient, we drop the subscript $_L$. We let $u, v, u', \ldots$ vary over *Symb*, $\varphi, \psi, \ldots$ over *Sent*, and $\Gamma, \Delta, \ldots$ over subsets sets of *Sent*. Also,

$$V_\varphi$$

is the set of symbols *occurring* in φ. Likewise, $V_\Gamma = \cup\{V_\varphi : \varphi \in \Gamma\}$.

2.2 Replacement

A *replacement* is a partial function ρ from *Symb* to *Symb* that respects categories: if $u \in dom(\rho)$ is of category $C \in Cat_L$, then so is $\rho(u)$. We write

$$\varphi[\rho]$$

for the result of replacing each occurrence of u in φ by $\rho(u)$.[7] It is convenient to assume that $V_\varphi \subseteq dom(\rho)$ — in words, ρ is a replacement *for* φ — so that ρ is the identity on symbols that don't get replaced.

We make the extra assumption that *Sent* is *closed* under replacement. Then the following conditions hold:[8]

(5)
a. If ρ is a replacement for φ, $\varphi[\rho] \in Sent$ and $V_{\varphi[\rho]} = range(\rho \restriction V_\varphi)$
b. $\varphi[id_{V_\varphi}] = \varphi$
c. If ρ, σ agree on V_φ, then $\varphi[\rho] = \varphi[\sigma]$.
d. $\varphi[\rho][\sigma] = \varphi[\sigma\rho]$, when σ is a replacement for $\varphi[\rho]$

2.3 Consequence relations

Following the Bolzano-Tarski tradition, we take consequence relations to hold between sets of sentences and sentences.[9]

Definition 2
A relation $R \subseteq \wp(Sent_L) \times Sent_L$ is

a. *reflexive* iff for all $\varphi \in Sent_L$, $\varphi\, R\, \varphi$;[10]
b. *transitive* iff whenever $\Delta\, R\, \varphi$ and $\Gamma\, R\, \psi$ for all $\psi \in \Delta$, we have $\Gamma\, R\, \varphi$;
c. *monotone* iff $\Delta\, R\, \varphi$ and $\Delta \subseteq \Gamma$ implies $\Gamma\, R\, \varphi$;
d. *truth-preserving* iff whenever $\Gamma\, R\, \varphi$ and (every sentence in) Γ is true, φ is also true.

[7] Other notations will also be used. For example, when $dom(\rho) = \{u_1, \ldots, u_n\}$ and $\rho(u_i) = u'_i$, we may write $\varphi[u_1/u'_1, \ldots, u_n/u'_n]$. Further, we often write just $\varphi[u/u']$ for the result applying a replacement ρ such that $\rho(u) = u'$ and ρ is the identity on all other symbols in φ.

[8] These are essentially the conditions in Peter Aczel's notion of a *replacement system* from Aczel (1990).

[9] More general notions take the conclusion to be a set of sentences as well, or use sequences or multisets instead of sets. In this paper we stick to the classical format.

[10] Writing $\psi\, R\, \varphi$ instead of $\{\psi\} R\, \varphi$.

A *consequence relation in* L is a reflexive, transitive, monotone, and truth-preserving relation (between sets of sentences and sentences). $CONS_L$ is the set of consequence relations in L. We let $\Rightarrow, \Rightarrow', \ldots$ vary over consequence relations.

If only finite sets of premises are considered, i.e. if $\Rightarrow \subseteq \wp^{<\omega}(Sent_L) \times Sent_L$, we say that $\Rightarrow$ is *finitary*. Results using the finiteness restriction will be marked (Fin). A weaker constraint is to consider *compact* consequence relations, in the sense that

If $\Gamma \Rightarrow \varphi$, then $\Gamma' \Rightarrow \varphi$ for some finite subset Γ' of Γ.

Define:

(6) a. $\Gamma \Rightarrow^{max} \varphi$ iff it is not the case that Γ is true and φ is false.

b. $\Gamma \Rightarrow^{min} \varphi$ iff $\varphi \in \Gamma$.

$\Rightarrow^{max}$ is essentially *material implication*. $\Rightarrow^{max}$ and $\Rightarrow^{min}$ are the smallest and the largest elements of the partial order $(CONS_L, \subseteq)$. Note also that for every truth-preserving relation R on $\wp(Sent_L) \times Sent_L$, there is a smallest consequence relation $cl_L(R)$ extending R. $cl_L(R)$ is the intersection of all consequence relations in which R is included.

3 Consequence from constants

Of particular interest are the consequence relations *generated* from a set of constants. The idea is familiar to every logician: φ follows from Γ, relative to a set X of constants, iff any reinterpretation of symbols *outside* X that makes the premises true also makes the conclusion true. In the Bolzano setting with an interpreted language we do not reinterpret symbols but *replace* them. As is well-known, this means that the availability of symbols in the language L may 'accidentally' effect the consequence relation. We will note such effects as we go along, and see what sort of assumptions about L will prevent them.

3.1 Bolzano consequence

Bolzano stressed the fact that we are in principle free to regard *any* set of symbols as constants. As pointed out in van Benthem (2003), one may thus think of Bolzano consequence as a *ternary* relation, between a set of premises, a conclusion, and a set X of symbols treated as constants. Equivalently, we shall define a *function* $\Rightarrow_{\_}$ from sets of symbols to consequence relations, as follows:[11]

[11] We do not follow Bolzano to the letter; for example, we do not require, as he did, that the set of premises should be *consistent* in order to have any consequences. For a discussion of this and several other aspects of Bolzano's notion of consequence, see van Benthem (2003). Another departure from Bolzano's original approach is that ours is syntactic, replacing *symbols*, whereas he replaced *concepts* ('Vorstellungen an sich'). Incidentally, this might make his account less vulnerable to detrimental effects due to lack of symbols in the language.

Definition 3
For any $X \subseteq Symb_L$, define the relation $\Rightarrow_X$ by

> $\Gamma \Rightarrow_X \varphi$ iff for every replacement ρ (for Γ and φ) that only moves symbols outside X (i.e. that is the identity on X), if $\Gamma[\rho]$ is true, so is $\varphi[\rho]$.

A relation of the form $\Rightarrow_X$ is called a *Bolzano consequence* (relation), and we let $BCONS_L$ be the set of Bolzano consequences in L.[12]

It is straightforward to verify the following claims.

Fact 4
(a) $BCONS_L \subseteq CONS_L$
(b) *In addition, Bolzano consequences are* base monotone, *in that*

$$X \subseteq Y \textit{ implies } \Rightarrow_X \subseteq \Rightarrow_Y$$

(c) $(BCONS_L, \subseteq)$ *is a partial order which has* $\Rightarrow_\emptyset$ *as its smallest and* $\Rightarrow_{Symb}$ *as its largest element.*

So $(BCONS_L, \subseteq)$ is a sub-order of $(CONS_L, \subseteq)$, and we see that

(7) $\quad \Rightarrow^{max} \; = \; \Rightarrow_{Symb}$

It often happens,[13] however, that

(8) $\quad \Rightarrow^{min} \; \subsetneq \; \Rightarrow_\emptyset$

so $BCONS_L$ can be a proper subset of $CONS_L$. The following is trivial but fundamental:

Lemma 5
(Replacement Lemma) *If* $\Gamma \Rightarrow_X \varphi$ *and* ρ *only replaces symbols outside* X*, then* $\Gamma[\rho] \Rightarrow_X \varphi[\rho]$.

Proof. Use the composition property (5d) in Section 2.2 of replacement, noting that if both ρ and σ only move symbols outside X, so does $\sigma\rho$. □

Furthermore, from base monotonicity and (5c) we see that only symbols occurring in premises and conclusion matter for Bolzano consequence:

Lemma 6
(Occurrence Lemma) $\Gamma \Rightarrow_X \varphi$ *if and only if* $\Gamma \Rightarrow_{X \cap V_{\Gamma \cup \{\varphi\}}} \varphi$.

[12] For readability, we use '$\Rightarrow_X$' in two ways: as a relation symbol, which enables us to write things like $\Gamma \Rightarrow_X \varphi$, and as the value of the function $\Rightarrow_{\_}\colon Symb_L \longrightarrow BCONS_L$ for the argument X.

[13] For example, in propositional logic, $***p \Rightarrow_\emptyset *p$, where p is a propositional letter and $*$ a unary truth function. Likewise, in all of the particular examples to follow, $\Rightarrow_\emptyset$ is distinct from $\Rightarrow^{min}$.

3.2 Example: propositional logic

Let PL be a standard language of *propositional logic*, whose symbols consist of a suitable set and an infinite supply of propositional letters — say, $Symb_{PL} = \{\neg, \vee, \wedge\} \cup \{p_0, p_1, \ldots\}$ with the (non-empty) categories 'unary truth function', 'binary truth function', and 'propositional letter' — and let $\models_{PL}$ be the corresponding (classical) consequence relation. The usual definition of consequence in this language is model-theoretic, but we can 'simulate' $\models_{PL}$ also in our substitutional setting, where $p_0, p_1, \ldots$ are sentences with fixed truth values, and the truth values of complex sentences are computed from these by the usual truth tables. Replacing proposition letters by others amounts to 'assigning' arbitrary truth values to them, under a simple assumption: let us say that PL, viewed as an interpreted language, is *non-trivial* iff the sequence of truth values of $p_0, p_1, \ldots$ is not eventually constant. Clearly,

(9) If PL is non-trivial, then for every (countable) sequence $\alpha_1, \alpha_2, \ldots$ of truth values there are propositional letters $p_{i_1}, p_{i_2}, \ldots$ such that the truth value of p_{i_j} is α_j, for all j.

Using (9), one easily verifies:

Fact 7
If PL is non-trivial, then $\Gamma \models_{PL} \varphi$ iff $\Gamma \Rightarrow_{\{\neg,\vee,\wedge\}} \varphi$.

3.3 Example: first-order logic

Now let FO be a standard language of *first-order logic*. We have $Symb_{FO} = \{\neg, \vee, \wedge, \forall, \exists, =\} \cup \{P_0, P_1, \ldots, c_0.c_1, \ldots\}$ with obvious (non-empty) categories such as 'type $\langle 1 \rangle$ quantifier', 'binary predicate symbol', 'individual constant', etc. We also assume that all symbols (also the non-logical ones) have fixed interpretations. Truth values of sentences are calculated as usual, and we let $\models_{FO}$ be the standard (classical) consequence relation.

At first blush, one might think the Bolzano approach simply amounts to FO with *substitutional interpretation of the quantifiers*, but this is not so. The reason is that in standard definitions of logical consequence with substitutional quantification, as in Dunn and Belnap (1968), only the quantifiers are interpreted substitutionally, but not the rest of the language. In more detail, in their substitutional account of FO, call it FO_{subst}, truth is defined relative to an arbitrary assignment ν of truth values to the atomic sentences, extended in the usual way to negations and conjunctions, and to universally quantified sentences by

(10) $\forall x \varphi(x)$ is true relative to ν iff $\varphi(c)$ is true relative to ν for all individual constants c.

Logical consequence is defined as follows:

(11) $\Gamma \models_{FO_{\text{subst}}} \varphi$ iff every assignment ν relative to which Γ is true is also one relative to which φ is true.

Dunn and Belnap show that if there are only finitely many individual constants, $c_1, \ldots, c_n$, then $\models_{FO_{\text{subst}}}$ will seriously diverge from $\models_{FO}$ since, for example,

$$\{Pc_1, \ldots, Pc_n\} \models_{FO_{\text{subst}}} \forall x Px,$$

but if there are infinitely many constants, *FO-validity* (consequence of the empty set) coincides with FO_{subst}-validity, although *FO-consequence* still differs from FO_{subst}-consequence, since we have

$$\{Pc\colon c \text{ is an individual constant}\} \models_{FO_{\text{subst}}} \forall x Px$$

In particular, $\models_{FO_{\text{subst}}}$ is not compact, in contrast with $\models_{FO}$.

In the Bolzano setting, on the other hand, there is no such thing as an arbitrary assignment of truth values to atomic sentences. These have fixed truth values, which can be 'varied' by replacing predicate symbols and individual constants to the extent that such are available. Every replacement corresponds to an assignment of truth values to atomic sentences, but the converse would be a rather substantial requirement. Without that, what we get is

(12) $\models_{FO} \; \subsetneq \; \models_{FO_{\text{subst}}} \; \subsetneq \; \Rightarrow_{\{\neg,\vee,\wedge,\forall,\exists,=\}}$

So $\models_{FO}$ is a consequence relation, but not in general a Bolzano consequence. However, we will see in Section 8 that in a slightly more general but still Bolzano style framework, $\models_{FO}$ is what we will call a *general* Bolzano consequence relation.

3.4 Two toy languages

We now describe in some detail two very simple languages and their consequence relations. These examples, and variants of them, will serve later on to illustrate various features of Bolzano consequence.

3.4.1 The language L_1

Let the language L_1 be specified as follows:

$$\begin{aligned} &\mathit{Symb}_{L_1} = \{R, a, b\} \quad \text{(with } a, b \text{ of the same category)} \\ &\mathit{Sent}_{L_1} = \{\mathit{Raa}, \mathit{Rab}, \mathit{Rba}, \mathit{Rbb}\} \\ &\mathit{Tr}_{L_1} = \{\mathit{Raa}, \mathit{Rab}, \mathit{Rbb}\} \end{aligned}$$

Here R is the only symbol of its category. It can only be replaced by itself, which means that it can in effect be disregarded. So in this and similar examples to follow, when writing things like $X \subseteq \mathit{Symb}_{L_1}$, we really mean $X \subseteq \mathit{Symb}_{L_1} - \{R\}$.

L_1 has the feature that no replacement of a single symbol can turn a true sentence into a false one; only a permutation of a and b can do that. We have, for example,

(13) $\emptyset \Rightarrow_\emptyset Raa$, $\emptyset \Rightarrow_\emptyset Rbb$, $\emptyset \not\Rightarrow_\emptyset Rab$, but if $a \in X$ or $b \in X$, then $\emptyset \Rightarrow_X Rab$

Note that the first claim already shows that $\Rightarrow_\emptyset \neq \Rightarrow^{min}$. As to the last claim of (13), since we are only allowed to replace symbols outside X, in this case at most one of a and b can be replaced, so the conclusion cannot be falsified. The following is a complete description of the Bolzano consequence relations in L_1:

Fact 8
In the language L_1:

(i) $\Rightarrow_\emptyset \ = \ cl_{L_1}(\{\langle\emptyset, Raa\rangle, \langle\emptyset, Rbb\rangle\})$

(ii) *If a or b belong to $X \subseteq Symb_{L_1}$, then*
$\Rightarrow_X \ = \ cl_{L_1}(\{\langle\emptyset, Raa\rangle, \langle\emptyset, Rbb\rangle, \langle\emptyset, Rab\rangle\})$

Proof. (i) By (13) and monotonicity, we need only consider inferences with the conclusion Rab. (I.e. for all $\Gamma \subseteq Sent_{L_1}$, $\Gamma \Rightarrow_\emptyset Raa$ follows from $\emptyset \Rightarrow_\emptyset Raa$ by monotonicity.) Suppose $\Gamma \Rightarrow_\emptyset Rab$. We can assume $Raa, Rbb \notin \Gamma$, by (13) and transitivity: if e.g. $Raa \in \Gamma$ then $\Gamma - \{Raa\} \Rightarrow_\emptyset Rab$. By reflexivity and monotonicity, we can also assume that $Rab \notin \Gamma$. But $Rba \not\Rightarrow_\emptyset Rab$ (permuting a and b makes the premise true and the conclusion false). So all valid inferences $\Gamma \Rightarrow_\emptyset \varphi$ in L_1 belong to the closure of the two listed in (i). The proof of (ii) is similar, using base monotonicity ($\Rightarrow_\emptyset \subseteq \Rightarrow_X$), and the last claim of (13). □

In particular,

(14) $\Rightarrow_\emptyset \subsetneq \Rightarrow_{\{b\}} = \Rightarrow_{\{a\}} = \Rightarrow_{\{a,b\}}$

3.4.2 The language L_2

L_2 just adds one symbol (of the same category) to L_1, but no new false sentences:

$$Symb_{L_2} = \{R, a, b, c\}$$
$$Sent_{L_2} = \{Rxy \colon x, y \in \{a, b, c\}\}$$
$$Tr_{L_2} = \{Raa, Rbb, Rcc, Rab, Rac, Rbc, Rca, Rcb\}$$

First, clearly,

(15) $\emptyset \Rightarrow_\emptyset Raa$, $\emptyset \Rightarrow_\emptyset Rbb$, $\emptyset \Rightarrow_\emptyset Rcc$, but if $x \neq y$, then $\emptyset \not\Rightarrow_\emptyset Rxy$

Next,

(16) If $x \neq y$ and $Rxy \notin \Gamma$, then $\Gamma \not\Rightarrow_\emptyset Rxy$.

For if ρ maps x to b, y to a, and the remaining symbol to c, it is a *permutation* of $Symb_{L_2}$, and then no sentence except Rxy is mapped to Rba, so all premises in Γ are true.

With respect to $\Rightarrow_{\{a\}}$, we have, in addition to the valid inferences with $\Rightarrow_{\emptyset}$,

(17) a. $\emptyset \Rightarrow_{\{a\}} Rab$ and $\emptyset \Rightarrow_{\{a\}} Rac$
b. $Rca \Rightarrow_{\{a\}} Rcb$ and $Rba \Rightarrow_{\{a\}} Rbc$

Next, if $Rba \notin \Gamma$, then Γ is true, and so cannot imply Rba. Also

(18) $\{Rba, Rbc, Rcb\} \not\Rightarrow_{\{a\}} Rca$ [map a to itself and permute b and c]

For $\Rightarrow_{\{a,c\}}$, the situation is quite simple, since the empty set now implies each sentence except Rba, and no set of premises not containing Rba implies Rba. Our findings can be summarized as follows:

Fact 9
Let Let $\Phi_0 = \{\langle\emptyset, Rxx\rangle : x \in Symb_{L_2}\}$, and $\Phi_1 = \Phi_0 \cup \{\langle\emptyset, Rab\rangle, \langle\emptyset, Rac\rangle\}$. In the language L_2:

(i) $\Rightarrow_{\emptyset} \; = \; cl_{L_2}(\Phi_0)$

(ii) $\Rightarrow_{\{a\}} \; = \; cl_{L_2}(\Phi_1 \cup \{\langle\{Rba\}, Rbc\rangle, \langle\{Rca\}, Rcb\rangle\})$

(iii) $\Rightarrow_{\{a,c\}} \; = \; cl_{L_2}(\{\langle\emptyset, Rxy\rangle : (x,y) \neq (b,a)\}) \; = \; \Rightarrow_{\{a,b,c\}}$

4 Minimality

L_1 and L_2 provide examples where different sets X, Y generate the same Bolzano consequence. One would expect sets that are *minimal* in this respect to be particularly well behaved. Perhaps the most obvious idea about minimality is the following.

Definition 10
X is *minimal* iff for all $u \in X$, $\Rightarrow_X \not\subseteq \Rightarrow_{X-\{u\}}$.

So X is minimal in the sense that if any one of its symbols is left out, a smaller consequence relation results. The other natural sense of 'minimal', as we noted, is minimality with respect to the sets generating the same consequence relation. In fact, it is easy to see that these two notions of minimality are equivalent.

$\emptyset$ is minimal in any language. In the languages L_1 and L_2 of Section 3.4, all singleton sets are minimal, since in each case, $\Rightarrow_{\{x\}}$ is distinct from $\Rightarrow_{\emptyset}$. Also, it is easy to see that $\{a, c\}$ is minimal in L_2, but $\{a, b\}$ in L_1, and $\{a, b, c\}$ in L_2, are not minimal.

Being minimal doesn't entail being the smallest set generating the same consequence relation; e.g. in L_1 there is a set of symbols X with distinct minimal subsets in $\{Y : \Rightarrow_Y = \Rightarrow_X\}$. However, in Westerståhl (2010) it was shown that there always exists at least one subset which is minimal among these, at least if we restrict attention to compact consequence relations (see Section 2.3):[14]

Theorem 11
For every $X \subseteq \mathit{Symb}_L$, if $\Rightarrow_X$ is compact, then X has a subset which is minimal among those generating $\Rightarrow_X$.

Thus, if we restrict attention to minimal subsets of *Symb*, no compact consequence relation of the form $\Rightarrow_X$ will be left out.

In this paper, we shall prove related and in a sense stronger results. First, we show that under some additional assumptions about the language L, there actually is a *smallest* subset generating $\Rightarrow_X$; moreover, this subset has a simple independent description (Corollary 26). Then we prove that the same result holds if we lift those restrictions (but not compactness), but use a slightly more general framework for Bolzano consequence (Corollary 39).

The requirement of compactness in Theorem 11, however, cannot be removed, as we now show.

Fact 12
There is a language L and a set $X \subseteq \mathit{Symb}_L$ of constants such that $\Rightarrow_X$ is not compact, and there is no minimal X' with $\Rightarrow_X = \Rightarrow_{X'}$.

Proof. We use a language $L_{\mathbb{N}}$ for arithmetic with numerals and predicates for any finite or co-finite set of natural numbers, plus a quantifier for "there are infinitely many". The *symbols* in $\mathit{Symb}_{L_{\mathbb{N}}}$ are thus constants c_n for every $n \in \mathbb{N}$, predicates P_A and $\neg P_A$ for every *finite* set A of numbers, and a predicate functor *Inf*. (So the non-empty categories here are 'numeral', '1-place predicate' and '1-place predicate functor'.) The *sentences* in $\mathit{Sent}_{L_{\mathbb{N}}}$ are of one of the forms $P_A c_n$, $\neg P_A c_n$, $\mathit{Inf}\, P_A$, and $\mathit{Inf} \neg P_A$. A sentence φ is in $\mathit{Tr}_{L_{\mathbb{N}}}$ iff φ is $P_A c_n$ and $n \in A$, or φ is $\neg P_A c_n$ and $n \notin A$, or φ is $\mathit{Inf} \neg P_A$. Note that $L_{\mathbb{N}}$ is countable.

Let $X = \{c_n\}_{n \in \mathbb{N}}$. In what follows, $(\neg)P_A$ stands for an arbitrary predicate P_A or $\neg P_A$. First:

(19) If there are infinitely many sentences of the form $(\neg)P_A c_i$ in Γ, then $\Gamma \Rightarrow_X \mathit{Inf}(\neg)P_A$.

This is because a replacement that makes all the sentences in Γ true then has to replace $(\neg)P_A$ by a predicate with an infinite extension (that is, a predicate of the form $\neg P_B$), since the c_i must not be replaced. But then, after such a replacement, the conclusion is also true. On the other hand,

(20) If $\mathit{Inf}(\neg)P_A \notin \Gamma$, and only finitely many sentences of the form $(\neg)P_A c_i$ are in Γ, then $\Gamma \not\Rightarrow_X \mathit{Inf}(\neg)P_A$.

[14] Actually, the proof in Westerståhl (2010) was given for finitary Bolzano consequence relations, but it is easily adapted to compact relations. That paper also identified a stricter notion of minimality, called *strong minimality*, and proved some results about it. These results are subsumed under the treatment in the present paper. In particular, in the more general Bolzano style framework introduced in Section 7, minimality and strong minimality coincide.

To see this, let A' be the finite set of numbers i such that $(\neg)P_A c_i \in \Gamma$. Consider the replacement ρ which replaces $(\neg)P_A$ by $P_{A'}$ and all other predicates by $\neg P_\emptyset$. Since $\mathit{Inf}(\neg)P_A \notin \Gamma$, it follows that all sentences in $\Gamma[\rho]$ are true (note that $\mathit{Inf}\neg P_\emptyset$ and all sentences $\neg P_\emptyset c_j$ are true), but $\varphi[\rho]$, i.e. $\mathit{Inf}\, P_{A'}$, is false. And since ρ does not act on X, this shows that $\Gamma \not\Rightarrow_X \varphi$. Next, we observe

(21) If $(\neg)P_A c_i \notin \Gamma$, then $\Gamma \not\Rightarrow_X (\neg)P_A c_i$.

For consider the replacement ρ which replaces $(\neg)P_A$ by $\neg P_{\{i\}}$ and all other predicates by $\neg P_\emptyset$. ρ doesn't act on X, all sentences in $\Gamma[\rho]$ are true (even if $\mathit{Inf}(\neg)P_A \in \Gamma$, since $\mathit{Inf}\neg P_{\{i\}}$ is true), but $(\neg)P_A c_i[\rho]$, i.e. $\neg P_{\{i\}} c_i$, is false.

This allows us to conclude:

(22) $\Rightarrow_X$ is not compact.

Take $\Gamma = \{P_{\{0\}} c_n\}_{n\in\mathbb{N}}$ and $\varphi = \mathit{Inf}\, P_{\{0\}}$. Then $\Gamma \Rightarrow_X \varphi$ by (19), but, by (20), there is no finite subset Γ' of Γ such that $\Gamma' \Rightarrow_X \varphi$.

Now let $X^- \subseteq X$ be any set of symbols such that the number of constants c_i which are in X but not in X^- is finite, and let $X^{--} \subseteq X$ be any set of symbols such that the number of constants c_i in X but not in X^{--} is infinite. Then we have:

(23) $\Rightarrow_{X^-} = \Rightarrow_X$

For suppose $\Gamma \Rightarrow_X \varphi$. We must show $\Gamma \Rightarrow_{X^-} \varphi$. This is clear if $\varphi \in \Gamma$, so suppose $\varphi \notin \Gamma$. It then follows from (21) that φ cannot be of the form $(\neg)P_A c_i$. So we have $\varphi = \mathit{Inf}(\neg)P_A$ for some A, and then it follows from (20) that there must be infinitely many sentences of the form $(\neg)P_A c_i$ in Γ. Thus, there are infinitely many sentences $(\neg)P_A c_i$ in Γ such that c_i is in X^-. So it is still the case that for a replacement ρ to make all the sentences in Γ true, ρ has to replace $(\neg)P_A$ by a predicate $\neg P_B$ with an infinite extension. Finally,

(24) $\Rightarrow_{X^{--}} \neq \Rightarrow_X$

Take for Γ all sentences of the form $P_{\{0\}} c_i$ for c_i not in X^{--}, and $\mathit{Inf}\, P_{\{0\}}$ for φ. By (19), $\Gamma \Rightarrow_X \varphi$, but now $\Gamma \not\Rightarrow_{X^{--}} \varphi$. Consider a replacement ρ such that $\rho(c_i) = c_0$ for all c_i not in X^{--}, but nothing else is moved. All sentences in $\Gamma[\rho]$ are true, since $\Gamma[\rho]$ is the singleton $\{P_{\{0\}} c_0\}$, but $\varphi[\rho]$, i.e. $\mathit{Inf}\, P_{\{0\}}$, is false.

Now the desired claim follows: there is no minimal subset X' of X such that $\Rightarrow_X = \Rightarrow_{X'}$. Subsets of X are either of the form X^- or X^{--}. But subsets of the form X^- are clearly not minimal, and subsets of the form X^{--} do not generate a consequence relation identical to $\Rightarrow_X$. $\square$

$$\begin{tikzcd}
(CONS_L, \subseteq) \arrow[rr, bend left, dotted, "C_{_}"] & & (\wp(Symb_L), \subseteq) \arrow[ll, bend left, dotted, "\Rightarrow_{_}"]
\end{tikzcd}$$

Figure 1: Logical consequence and constant extraction

5 Extracting constants from consequence relations

5.1 Defining extraction

We now introduce an operation corresponding to the extraction of logical constants from a consequence relation. When a particular consequence relation is given, certain symbols are to be considered as logical constants because the consequence relation makes them play a special role with respect to validity. As explained in the Introduction, our guiding intuition is that a symbol is constant if replacing it can destroy *at least one* inference.

Definition 13
The function $C_{\_}\colon CONS_L \to \wp(Symb_L)$ is defined for $\Rightarrow \in CONS_L$ by:

$$u \in C_{\Rightarrow} \text{ iff there are } \Gamma,\ \varphi, \text{ and } u' \text{ such that } \Gamma \Rightarrow \varphi \text{ but } \Gamma[u/u'] \not\Rightarrow \varphi[u/u']$$

We first observe, as a direct consequence of the Replacement Lemma, that when $C_{\_}$ is applied to a Bolzano consequence relation, it will never pick out a non-logical constant:

Fact 14
For all $X \subseteq Symb$, $C_{\Rightarrow_X} \subseteq X$.

Proof. Suppose $u \in C_{\Rightarrow_X}$, and Γ, φ, and u' are as above. If $u \notin X$ we would have $\Gamma[u/u'] \Rightarrow_X \varphi[u/u']$ by Replacement. So $u \in X$. □

Logical consequence can be construed as a function from sets of symbols to consequence relations. Extraction goes in the opposite direction. Moreover, the domains of both functions are naturally ordered by inclusion, so the situation is as shown in Figure 1. Fact 4(b) said that $\Rightarrow_{\_}$ is an order-preserving mapping from $(\wp(Symb_L), \subseteq)$ to $(CONS_L, \subseteq)$. We would like $C_{\_}$ to provide some sort of inverse order-preserving mapping. Before looking into this and other properties of $C_{\_}$, let us see some examples of how $C_{\_}$ works.

5.2 Examples

There is one case when the function $C_{\_}$ trivially fails to yield the intended result because of its substitutional character, namely, when a symbol u is unique in its category. Then there is no other symbol to replace u with, so it will not count as a logical constant, no matter what inferential role it plays. This situation arises with negation, which is usually the only unary connective in logical languages. To sidestep this difficulty, we shall assume, when considering propositional logic or first-order logic, that they come equipped with another unary connective, say $\dagger$, interpreted by the constant unary truth-function 'equal to false'.[15] With this assumption, we can verify that $C_{\_}$ satisfies the first criterion mentioned in the Introduction for a reasonable 'extraction function': it gives the correct set of logical constants in familiar logical languages.

5.2.1 Familiar logical languages

Fact 15
$C_{\models_{PL}}$ is the standard set of logical constants of propositional logic.

Proof. $p \models_{PL} p \vee q$ but $p \not\models_{PL} p \wedge q$. That is, replacing $\vee$ by $\wedge$ destroys the validity of the first inference, so $\vee \in C_{\models_{PL}}$. Likewise, $p \models_{PL} \neg\neg p$ but $p \not\models_{PL} \dagger\dagger p$, and thus $\neg \in C_{\models_{PL}}$. Similarly for other familiar constants. On the other hand, (uniformly) replacing propositional letters can never destroy a valid $\models_{PL}$-inference. □

Recall from Section 3.2 that if the interpreted propositional language is non-trivial (the sequence of truth values of $p_0, p_1, \ldots$ is not eventually constant), then $\models_{PL}$ is a Bolzano consequence, say, $\models_{PL} = \Rightarrow_{\{\neg,\dagger,\vee,\wedge\}}$. But the fact that $C_{\_}$ recovers the right constants doesn't depend on this. We get, with the same kind of argument as above, the correct result also for first-order logic, even though neither $\models_{FO}$ nor $\models_{FO_{\mathrm{subst}}}$ (Section 3.3) is usually a Bolzano consequence relation:

Fact 16
$C_{\models_{FO}} = C_{\models_{FO_{subst}}}$ is the standard set of logical constants of first-order logic.

Indeed, most familiar consequence relations are such that suitably replacing a logical symbol can destroy an inference, while this is not possible for non-logical symbols. For example, consider an intuitionistic propositional logic whose consequence relation is defined not model-theoretically but axiomatically (with suitable extra axioms for $\dagger$). Again, virtually the same kind of arguments show that $C_{\_}$ extracts precisely the logical symbols from this relation.

5.2.2 Application to L_1 and L_2

$C_{\_}$ behaves rather badly for L_1, since $C_{\Rightarrow_X} = \emptyset$ for all $X \subseteq Symb_{L_1}$. This is because replacing just *one* symbol can never destroy a $\Rightarrow_X$-inference in L_1; you need to replace two symbols simultaneously. Can we revise the extraction

[15] An alternative option is to simply recognize that in a substitutional framework, the question whether some symbol which is alone in its category is constant or not, is deprived of interest.

method so that it handles such cases better? There is actually a way (see Section 9.3), but in in this paper we shall stick to the function $C\_$ and make it behave better by placing requirements on the language. For now, we observe that already in L_2, which contains just one extra symbol (of the same category), the situation is significantly different.

First, note that replacing a by c destroys the inference $\Rightarrow_{\{a\}} Rab$.[16] Thus, $a \in C_{\Rightarrow_{\{a\}}}$, so by Fact 14,

(25) $\quad C_{\Rightarrow_{\{a\}}} = \{a\}$

Next, since $\Rightarrow_{\{a,c\}} Rbc$ but $\not\Rightarrow_{\{a,c\}} Rba$, $c \in C_{\Rightarrow_{\{a,c\}}}$ in L_2. But $a \notin C_{\Rightarrow_{\{a,c\}}}$; this follows by checking that replacing a by b or c does not destroy any of the basic inferences listed in Fact 9 (iii). Thus, in L_2,

(26) $\quad C_{\Rightarrow_{\{a,c\}}} = \{c\}$

So the situation for L_2 is better than for L_1, but it is still not good, at least if we want $C\_$ to be an order-preserving inverse on $CONS_L$. The failure of order preservation is no surprise given that there are both a positive and a negative condition in the definition of $C\_$. The witness to a non-valid inference might disappear by shifting to a bigger consequence relation. Perhaps more surprisingly, the situation is no better for Bolzano consequences.

Fact 17
There are languages L and sets $X, Y \subseteq Symb_L$ such that:

(a) $\Rightarrow_X \subseteq \Rightarrow_Y$ *but* $C_{\Rightarrow_X} \not\subseteq C_{\Rightarrow_Y}$

(b) $\Rightarrow_X \not\subseteq \Rightarrow_{C_{\Rightarrow_X}}$

Proof. An example is provided by (25) and (26) for L_2. There we have $\Rightarrow_{\{a\}} \subseteq \Rightarrow_{\{a,c\}}$ by base monotonicity, but $\{a\} = C_{\Rightarrow_{\{a\}}} \not\subseteq C_{\Rightarrow_{\{a,c\}}} = \{c\}$. Also, $\Rightarrow_{\{a,c\}} \not\subseteq \Rightarrow_{C_{\Rightarrow_{\{a,c\}}}} = \Rightarrow_{\{c\}}$, since, for example, $\emptyset \Rightarrow_{\{a,c\}} Rab$ but $\emptyset \not\Rightarrow_{\{c\}} Rab$. $\square$

6 A Galois connection under special assumptions

Fact 17 shows that $C\_$ does not yet behave in the way we would like. On the other hand, $C\_$ passes the first part of the test: it gives the right results when applied to familiar logical systems. Our diagnosis will be that the problems are due to particular features of the languages used in the counter-examples, rather to shortcomings of the definition itself. The present section isolates a subclass of languages, sets of constants, and consequence relations for which $C\_$ behaves well, within the classical Bolzano framework introduced in Section 2. In the

[16] Here and in what follows we write $\Rightarrow_X \varphi$ rather than $\emptyset \Rightarrow_X \varphi$; meaning that φ is *valid* (relative to $\Rightarrow_X$).

next section we will see that with a slight extension of that framework, many (but not all) of those restrictions can be lifted.

6.1 A factorization property for replacements

Let us take a closer look at the failure of monotonicity with respect to Bolzano consequence relations. In L_2, as we saw, $\Rightarrow_{\{a\}} \subseteq \Rightarrow_{\{a,c\}}$ but $C_{\Rightarrow_{\{a\}}} \not\subseteq C_{\Rightarrow_{\{a,c\}}}$, the reason being that $a \in C_{\Rightarrow_{\{a\}}}$ but $a \notin C_{\Rightarrow_{\{a,c\}}}$.

Relative to $\Rightarrow_{\{a\}}$ (and to $\Rightarrow_{\{a,c\}}$ as well), a *should* clearly be identified as a constant. After all, holding a fixed does make a difference, e.g. $\Rightarrow_{\{a\}} Rab$ but $\not\Rightarrow_{\emptyset} Rab$ (recall that Rba is false). But this is not sufficient for $C_{\_}$ to spot a as a constant. $\Rightarrow_{\{a\}} Rab$ and $\not\Rightarrow_{\{a\}} Rba$, but Rab cannot be turned into Rba by replacing only a, as the definition of $C_{\_}$ requires. For $\Rightarrow_{\{a\}}$, this is not a problem, because the non-constant symbol c can be used as a stop-over on the journey. Instead of jumping from the validity of Rab to the falsity of Rba, one can stop by the invalidity of Rcb. Then $a \in C_{\Rightarrow_{\{a\}}}$, because $\Rightarrow_{\{a\}} Rab$ and $\not\Rightarrow_{\{a\}} Rab[a/c]$.

Shifting to $\Rightarrow_{\{a,c\}}$, things are different: $\Rightarrow_{\{a,c\}} Rab[a/c]$. As it happens, there is no alternative way in L_2 to witness the constancy of a, and a ends up being outside $C_{\Rightarrow_{\{a,c\}}}$. But consider a language L_3 which is just as L_2 except that it contains another symbol d of the same category as c. So $Tr_{L_2} = Tr_{L_3} \cap Sent_{L_2}$, and let us also assume that Rad is true. In L_3, the situation improves because d can be used as a substitute stop-over: now $a \in C_{\Rightarrow_{\{a,c\}}}$, because $\Rightarrow_{\{a,c\}} Rab$ and $\not\Rightarrow_{\{a,c\}} Rab[a/d]$.

The lesson we would like to draw is that monotonicity holds when the language is rich enough (so that a d is available) and fails when it is not. We shall now spell out a general factorization property for replacements, which reflects the availability of stop-overs, and connect it to properties of languages, sets of constants, and consequence relations. This will enable us to prove monotonicity and more.

Definition 18 (Factorization Property)
Let $X, Y \subseteq Symb_L$ and $\Delta \subseteq Sent_L$. We say that *X-replacements in Δ factor through Y* iff for any replacement ρ which is defined on V_Δ and acts outside X, there are replacements σ and τ such that:

(i) σ acts only on $Y - X$
(ii) $\sigma(Y - X) \cap V_\Delta = \emptyset$
(iii) τ acts outside Y
(iv) $\rho = \tau \circ \sigma$

where a replacement *acts outside a set* if it is the identity on every element in that set for which it is defined, and *acts on a set* if every element for which it is not the identity is in that set.

Note that σ needs to have the same domain as ρ, whereas the domain of τ can be taken to be no wider than the range of σ. Sentences in Δ will typically be the sentences in an inference $\Gamma \Rightarrow_X \varphi$. $\sigma(Y - X) \cap V_\Delta = \emptyset$ then means that σ replaces symbols in $Y - X$ by 'new' symbols not occurring in the inference. In our earlier example, with $X = \{a\}$ and $Y = \{a, c\}$, σ corresponds to replacing c with d, so that everything that a replacement ρ which moves c could do can now be done by a replacement τ moving d instead of c.

When is this factorization possible, i.e. when are helpful symbols like d available? First, d qualified as a substitute for c because it was of the same category as c and did not belong to the old set of constants. In order to secure availability of such ds, a simple requirement would be that there are infinitely many symbols in each non-empty category. In Bonnay and Westerståhl (2010) we called such languages *rich*. But we also need that these rich resources cannot be all consumed by the chosen constants. For each non-empty category, there should always be infinitely many symbols in that category not taken as constants. The simplest requirement would be to restrict attention to *finite* sets of symbols. But we can replace these two by the single weaker requirement that we only consider *co-infinite* sets of symbols, i.e. in each non-empty category, there are infinitely many symbols in $Symb_L - X$. Let $\wp^{\text{coinf}}(Symb_L)$ be the set of such sets of symbols. As long as (in each non-empty category) $\wp^{\text{coinf}}(Symb_L)$ is not empty, assuming that the sets of symbols discussed are co-infinite entails assuming that L is rich.

In addition we need an assumption on the consequence relations. For simplicity, we shall assume that they are finitary, i.e. that only finite sets of premises are considered — marked by writing (FIN) — but in fact our proofs work under the weaker hypothesis that they are *compact* (Section 2.3; we shall indicate the required changes in the proofs).[17]

Co-infiniteness of the set of constants and finiteness of the set of sentences guarantee that new symbols are available, so the factorization property holds:

Lemma 19
If $Y \in \wp^{coinf}(Symb_L)$ and Δ is a finite set of L-sentences, then for all $X \subseteq Symb_L$, X-replacements in Δ factor through Y.

Proof. Since Y is co-infinite in each non-empty category, so is $Y - X$. Since moreover Δ is finite, for every symbol a_i in $(Y - X) \cap V_\Delta$, there is a different symbol b_i which is of the same category as a_i but does not belong to V_Δ or to Y. Define σ by

$$\sigma(x) = \begin{cases} b_i & \text{if } x = a_i \\ x & \text{otherwise} \end{cases}$$

Then define τ on the range of σ by

[17]Rather than (FIN), we could instead use the actually weaker requirement that the set of symbols occuring in an inference $\Gamma \Rightarrow \varphi$ is finite. Interestingly, our proofs do not go through if the set of symbols occurring in $\Gamma \cup \{\varphi\}$ is assumed to be co-infinite (in each non-empty category). This asymmetry suggests that the finiteness requirement does not play the same role for sets of symbols as it does for sets of premises. This is one reason we choose to work with the more precise if less simple assumption of co-infinity, rather than with richness and finite sets of symbols.

$$\tau(x) = \begin{cases} \rho(a_i) & \text{if } x = b_i \\ \rho(x) & \text{otherwise} \end{cases}$$

It is easy to check that $\rho = \tau \circ \sigma$ and all other conditions in Definition 18 are satisfied. $\square$

6.2 Monotonicity and preservation

The factorization property ensures monotonicity of $C_{\_}$ with respect to Bolzano consequence relations. The proof hinges on the same kind of reasoning we went through in the example.

Theorem 20
(Fin) *If Y is co-infinite, then $\Rightarrow_X \subseteq \Rightarrow_Y$ implies $C_{\Rightarrow_X} \subseteq C_{\Rightarrow_Y}$.*

Proof. Assume $\Rightarrow_X \subseteq \Rightarrow_Y$, where Y is co-infinite, and $u \in C_{\Rightarrow_X}$. We want to show that $u \in C_{\Rightarrow_Y}$. By definition of $C_{\_}$, there are Γ, φ and u' in L such that $\Gamma \Rightarrow_X \varphi$ and $\Gamma[u/u'] \not\Rightarrow_X \varphi[u/u']$. By definition of $\Rightarrow_X$, there is a replacement ρ acting outside of X such that the sentences in $\Gamma[u/u'][\rho]$ are true but $\varphi[u/u'][\rho]$ is false.

The hypotheses of Lemma 19 apply with respect to X, Y, and $\Delta = \Gamma \cup \{\phi\} \cup \Gamma[u/u'] \cup \{\phi[u/u']\}$, because of (Fin).[18] So X-replacements in Δ factor through Y, i.e. there are σ and τ such that σ acts only on $Y - X$, $\sigma(Y - X) \cap (V_\Delta) = \emptyset$, τ acts outside Y, and $\rho = \tau \circ \sigma$.

Since σ acts outside X we get, by Replacement,

$$\Gamma[\sigma] \Rightarrow_X \varphi[\sigma]$$

Hence, by assumption,

$$\Gamma[\sigma] \Rightarrow_Y \varphi[\sigma]$$

It is now sufficient to prove

$$\Gamma[\sigma][u/\sigma(u')] \not\Rightarrow_Y \varphi[\sigma][u/\sigma(u')] \tag{27}$$

Since $\sigma(Y - X) \cap V_\Delta = \emptyset$, it follows that $[\sigma][u/\sigma(u')] = [u/u'][\sigma]$. Hence, since $\rho = \tau \circ \sigma$, $\Gamma[\sigma][u/\sigma(u')][\tau]$ is $\Gamma[\rho]$, a set of true sentences, and $\varphi[\sigma][u/\sigma(u')][\tau]$ is $\varphi[\rho]$, a false sentence. Since τ acts outside Y, this proves (27). $\square$

In a similar manner, we can establish a more satisfactory inverse relationship between the mappings $\Rightarrow_{\_}$ and $C_{\_}$ restricted to Bolzano consequences. The consequence relation generated by *any* co-infinite set X of symbols is the same as what you get by first extracting the constants from $\Rightarrow_X$ and then generating

[18] If Γ were infinite but $\Rightarrow_X$ compact, the remainder of the proof would go through working with a finite subset Γ' of Γ such that $\Gamma' \Rightarrow_X \varphi$ and $\Gamma'[u/u'] \not\Rightarrow_X \varphi[u/u']$. Similarly for the proof of Theorem 21 below.

the Bolzano consequence from those constants, even if they form a proper subset of X.

Theorem 21
(FIN) *If X is co-infinite, $\Rightarrow_X = \Rightarrow_{C_{\Rightarrow_X}}$*

Proof. $C_{\Rightarrow_X} \subseteq X$ already implies $\Rightarrow_{C_{\Rightarrow_X}} \subseteq \Rightarrow_X$, so all we need to prove is $\Rightarrow_X \subseteq \Rightarrow_{C_{\Rightarrow_X}}$. Assume $\Gamma \Rightarrow_X \varphi$. We must show $\Gamma \Rightarrow_{C_{\Rightarrow_X}} \varphi$. Let ρ be a replacement acting outside of $C_{\Rightarrow_X}$. It is sufficient to show

(28) $\quad \Gamma[\rho] \Rightarrow_X \phi[\rho]$

The hypotheses in Lemma 19 apply, since X is co-infinite and $\Delta = \Gamma \cup \{\varphi\}$ is finite. So $C_{\Rightarrow_X}$-replacements in $\Gamma \cup \{\phi\}$ factor through X. Hence we have σ and τ such that σ acts only on $X - C_{\Rightarrow_X}$, $\sigma(X - C_{\Rightarrow_X}) \cap V_\Delta = \emptyset$, τ acts outside X, and $\rho = \tau \circ \sigma$.

First, we show that

(29) $\quad \Gamma[\sigma] \Rightarrow_X \varphi[\sigma]$

Since σ can be taken to be defined on the finite vocabulary of $\Gamma \cup \{\varphi\}$, σ acts on $X - C_{\Rightarrow_X}$, and no symbol in that set is replaced by a symbol occurring in $\Gamma \cup \{\varphi\}$, we have $\sigma = \sigma_n$ for some n, where, for $i \leq n$,

$$\sigma_i = id \cup \{(a_1, \sigma(a_1)), \ldots, (a_i, \sigma(a_i))\}$$

for some $a_1, \ldots, a_n \in X - C_{\Rightarrow_X}$, and id is the identity function on the rest of $V_{\Gamma \cup \{\varphi\}}$. Moreover, by the choice of the $\sigma(a_i)$, replacing $a_1, \ldots, a_n$ simultaneously according to σ and successively replacing them one by one gives the same result: for $\psi \in \Gamma \cup \{\varphi\}$,

$$\psi[\sigma_{i+1}] = \psi[\sigma_i][a_{i+1}/\sigma(a_{i+1})]$$

Assume for contradiction that a_{i+1} is the first symbol in the sequence for which consequence is not preserved, that is $\Gamma[\sigma_{i+1}] \not\Rightarrow_X \varphi[\sigma_{i+1}]$, but $\Gamma[\sigma_i] \Rightarrow_X \varphi[\sigma_i]$. So $\Gamma[\sigma_i][a_{i+1}/\sigma(a_{i+1})] \not\Rightarrow_X \varphi[\sigma_i][a_{i+1}/\sigma(a_{i+1})]$, but then $a_{i+1} \in C_{\Rightarrow_X}$, a contradiction. This proves (29).

Second, by Replacement, since τ acts outside X,

$$\Gamma[\sigma][\tau] \Rightarrow_X \varphi[\sigma][\tau]$$

Since $\rho = \tau \circ \sigma$, this proves (28). □

Theorem 21 relies on two assumptions: that X is co-infinite and that we consider only finitary (or compact) consequence relations. That none of these assumptions can be dropped follows from the next two facts.

Fact 22
(FIN) *There is a language L and a set $X \subseteq Symb_L$ with finite complement such that $\Rightarrow_X \not\subseteq \Rightarrow_{C_{\Rightarrow_X}}$.*

Proof. Consider the language L_2', which is a rich variant of L_2: $Symb_{L_2'} = \{R, a, b, c_0, c_1, \ldots\}$, $Sent_{L_2'} = \{Rxy : x, y \in Symb_{L_2'}\}$, and $Tr_{L_2'} = Sent_{L_2'} - \{Rba\}$. Now let

$$X = \{a, c_0, c_1, ...\}$$

Then we claim

$$a \notin C_{\Rightarrow_X}$$

Otherwise, there would be a finite set Γ, and φ, u' such that $\Gamma \Rightarrow_X \varphi$ but $\Gamma[a/u'] \not\Rightarrow_X \varphi[a/u']$. The latter means that there would be a replacement of b only — since b is the only symbol outside X — such that $\Gamma[a/u'][b/b']$ is true and $\varphi[a/u'][b/b']$ is false. Now observe that $\varphi[a/u']$ does not contain a; otherwise $u' = a$ which contradicts the assumptions. If $b' = b$, then $\varphi[a/u'][b/b']$ does not contain a either. If $b' \neq b$, then $\varphi[a/u'][b/b']$ does not contain b. So $\varphi[a/u'][b/b']$ is a sentence of the form Rxy which does not contain both a and b. But all those sentences are true: contradiction.

Next, note that $\Rightarrow_X Rc_i a$ but $\not\Rightarrow_X Rba$. This shows that each c_i belongs to $C_{\Rightarrow_X}$, and so

(30) $\quad C_{\Rightarrow_X} = X - \{a\} = \{c_0, c_1, \ldots\}$

But now observe that $\Rightarrow_X Rab$ but $\not\Rightarrow_{X-\{a\}} Rab$. Together with (30) this proves $\Rightarrow_X \not\subseteq \Rightarrow_{C_{\Rightarrow_X}}$. □

Fact 23
There is a language L and a co-infinite set $X \subseteq Symb_L$ such that $\Rightarrow_X$ is not compact and $\Rightarrow_X \not\subseteq \Rightarrow_{C_{\Rightarrow_X}}$.

Proof. We use a variant $L_{\mathbb{N}}'$ of the arithmetical toy language $L_{\mathbb{N}}$ from the proof of Lemma 12 in Section 4. To the symbols of $L_{\mathbb{N}}$ we add predicate functors Inf_n for $n \geq 1$, letting $Inf_0 = Inf$. We also add new numerals z_n for $n \geq 0$. The predicate symbols are the same, and the intuitive idea is that each Inf_n means the same as Inf (i.e. 'is infinite'), c_i denotes the number i as before, and z_n may denote any number. The sentences have the same forms as in $L_{\mathbb{N}}$, using also the new predicate functors and numerals. Thus, for each finite $A \subseteq \mathbb{N}$, $Inf_n \neg P_A$ is true, $Inf_n P_A$ is false, and if d is a numeral denoting i, then $P_A d$ is true iff $i \in A$, and $\neg P_A d$ is true iff $i \notin A$.

As before, let $X = \{c_0, c_1, \ldots\}$. Note that in $L_{\mathbb{N}}'$, X is co-infinite (in each non-empty category). Now claims corresponding to (19) – (24) in the proof of Lemma 12 go through in $L_{\mathbb{N}}'$ as well. First, with the same proof, we have

(31) $\quad$ If there are infinitely many sentences of the form $(\neg)P_A c_i$ in Γ, then $\Gamma \Rightarrow_X Inf_n(\neg)P_A$.

Likewise, with very small changes, we obtain

(32) If Γ contains no sentence of the form $\mathit{Inf}_m(\neg)P_A$, and only finitely many sentences of the form $(\neg)P_A c_i$, then $\Gamma \not\Rightarrow_X \mathit{Inf}_n(\neg)P_A$.

Also, for any numeral d and finite set A,

(33) If $(\neg)P_A d \notin \Gamma$, then $\Gamma \not\Rightarrow_X (\neg)P_A d$.

For suppose d denotes the number i, and consider a replacement ρ mapping $(\neg)P_A$ to $\neg P_{\{i\}}$ and all other predicate symbols to $\neg P_\emptyset$, while Inf_n is mapped to *Inf*, d is mapped to c_i (so if d *is* c_i, it is not moved), all the other z_k are mapped to some c_j with $j \neq i$, and no c_k is moved. Since $(\neg)P_A d \notin \Gamma$, one readily checks that $\Gamma[\rho]$ is true but $\varphi[\rho] = \neg P_{\{i\}} c_i$ is false.

Now, we claim:

(34) $C_{\Rightarrow_X} = \emptyset$

It suffices to show that $c_i \notin C_{\Rightarrow_X}$, for all i. Otherwise, there are i, Γ, φ, and d such that $\Gamma \Rightarrow_X \varphi$ but $\Gamma[c_i/d] \not\Rightarrow_X \varphi[c_i/d]$. It follows that $\varphi \notin \Gamma$, and thus by (33), that φ is not of the form $(\neg)P_A d'$. So φ has to be $\mathit{Inf}_n(\neg)P_A$ for some A and n. Since $\mathit{Inf}_n(\neg)P_A$ is not in Γ, hence not in $\Gamma[c_i/d]$ either, no sentence of the form $\mathit{Inf}_m(\neg)P_A$ is in $\Gamma[c_i/d]$; this follows since obviously, for all m, n,

$$\mathit{Inf}_m(\neg)P_A \Rightarrow_X \mathit{Inf}_n(\neg)P_A$$

So no sentence of the form $\mathit{Inf}_m(\neg)P_A$ is in Γ, and then (32) implies that there are infinitely many sentences of the form $(\neg)P_A c_j$ in Γ, since $\Gamma \Rightarrow_X \varphi$. However, from (31) it follows that only finitely many sentences of this form belong to $\Gamma[c_i/d]$, since $\Gamma[c_i/d] \not\Rightarrow_X \varphi$. But this is impossible, since $\Gamma[c_i/d]$ results from Γ by replacing c_i in at most one such sentence. This proves (34).

The example $\Gamma = \{P_{\{0\}} c_n : n \in \mathbb{N}\}$ and $\varphi = \mathit{Inf}\, P_{\{0\}}$ shows as before that $\Rightarrow_X$ is not compact. It also shows that $\Rightarrow_X \not\subseteq \Rightarrow_{C_{\Rightarrow_X}} = \Rightarrow_\emptyset$, since $\Gamma \Rightarrow_X \varphi$, but, clearly, $\Gamma \not\Rightarrow_\emptyset \varphi$. □

6.3 A Galois connection

Let us take stock. What kind of correspondence do we get between $C_{\_}$ and $\Rightarrow_{\_}$? We wanted something as close as possible to an isomorphism, with as few assumptions as possible. A relevant notion of correspondence in that context is the notion of a *Galois connection*. A Galois connection is a quadruple $\langle \mathcal{A}, \mathcal{B}, f, g\rangle$ with $\mathcal{A}$ and $\mathcal{B}$ two ordered structures, $f\colon A \to B$ and $g\colon B \to A$ two functions, such that the following four conditions hold: (I) f is monotone, (II) g is monotone, (III) $g \circ f$ is increasing, (IV) $f \circ g$ is decreasing.[19] An interesting

[19] A more compact characterization is that for all $a \in A$ and $b \in B$,

$$a \leq g(b) \text{ iff } f(a) \leq b$$

This is equivalent to the combination of (I)–(IV).

property of Galois connections is that f and g, even though they do not constitute a full-blown isomorphism, *give rise to an isomorphism.* From (I)–(IV), one can prove that f is an isomorphism with inverse g between the well-behaved subsets $g(B)$ and $f(A)$ of A and B.

Consider now an arbitrary language L. Think of $(\mathit{CONS}_L, \subseteq)$, the set of all consequence relations on L ordered by inclusion, as $\mathcal{A}$, and $(\wp(\mathit{Symb}_L), \subseteq)$, the set of all possible sets of constants ordered by inclusion, as $\mathcal{B}$. $C_{\_}$ and $\Rightarrow_{\_}$ are candidates for providing a Galois connection between $\mathcal{A}$ and $\mathcal{B}$. Base monotonicity (Fact 4) says that $\Rightarrow_{\_}$ is monotone — this is condition (II). The fact that the set of constants extracted from a Bolzano consequence relation $\Rightarrow_X$ is included in the original set X of constants (Fact 14) says that $C_{\Rightarrow_{\_}}$ is decreasing — this is condition (IV). Conditions (I) and (III) do not hold in general, not even (Fact 17) when attention is restricted from CONS_L to the proper subset BCONS_L of Bolzano consequence relations. However, for BCONS_L, suitable assumptions give us what we need: Theorem 20 is condition (I), and Theorem 21 implies that $\Rightarrow_{C_{\_}}$ is increasing, this is condition (III). Putting this together, we get

Theorem 24
(Fin) *$C_{\_}$ and $\Rightarrow_{\_}$ constitute a Galois connection between $(\mathit{BCONS}_L^{\mathrm{coinf}}, \subseteq)$ and $(\wp^{\mathrm{coinf}}(\mathit{Symb}_L), \subseteq)$.*

Here $\mathit{BCONS}_L^{\mathrm{coinf}}$ is the set of consequence relations of the form $\Rightarrow_X$ for some $X \in \wp^{\mathrm{coinf}}(\mathit{Symb}_L)$.

Our Galois connection is rather special in that the image of $\wp^{\mathrm{coinf}}(\mathit{Symb}_L)$ under $\Rightarrow_{\_}$ is the whole of $\mathit{BCONS}_L^{\mathrm{coinf}}$.[20] This reflects the fact that all of *CONS* or *BCONS* could not be part of the connection: restriction to the image of $\wp^{\mathrm{coinf}}(\mathit{Symb}_L)$ under $\Rightarrow_{\_}$ is needed not only to get an isomorphism but already to satisfy conditions (I) and (III). Indeed, we do not have a characterization of the action of $C_{\_}$ on consequence relations which are not of the form $\Rightarrow_X$ for some X (but see the informal discussion in Section 9.2).

Of special interest is now the image of $\mathit{BCONS}_L^{\mathrm{coinf}}$ under $C_{\_}$. Which well-behaved subset of $\wp^{\mathrm{coinf}}(\mathit{Symb}_L)$ gets selected by the Galois connection to be the codomain of the isomorphism? The answer is given by the next result.

Corollary 25
(Fin) *The image under $C_{\_}$ of $\mathit{BCONS}_L^{\mathrm{coinf}}$ is the set of minimal sets in $\wp^{\mathrm{coinf}}(\mathit{Symb}_L)$.*

Proof. First, to prove that every minimal co-infinite set X is the image of some $\Rightarrow_Y$ under $C_{\_}$, we prove that it is the image of the consequence relation generated by itself, that is:

If X is minimal and co-infinite, $X = C_{\Rightarrow_X}$.

[20] This also corresponds to the fact that not only is $g \circ f$ increasing, but actually $g \circ f = Id_A$ as stated in Theorem 21.

Because of Fact 14, we need only show $X \subseteq C_{\Rightarrow_X}$. This follows from $\Rightarrow_X = \Rightarrow_{C_{\Rightarrow_X}}$ (Theorem 21), since X is minimal and co-infinite.

Second, we prove:

For every co-infinite X, $C_{\Rightarrow_X}$ is minimal.

Take $u \in C_{\Rightarrow_X}$. We must show that $\Rightarrow_{C_{\Rightarrow_X}} \not\subseteq \Rightarrow_{C_{\Rightarrow_X}-\{u\}}$. By definition, there are Γ, φ, and u' such that $\Gamma \Rightarrow_X \varphi$ but $\Gamma[u/u'] \not\Rightarrow_X \varphi[u/u']$. So, by Theorem 21, $\Gamma \Rightarrow_{C_{\Rightarrow_X}} \varphi$. Also, by Replacement, $\Gamma \not\Rightarrow_{X-\{u\}} \varphi$. Since $C_{\Rightarrow_X} - \{u\} \subseteq X - \{u\}$ we get, by base monotonicity, $\Gamma \not\Rightarrow_{C_{\Rightarrow_X}-\{u\}} \varphi$. □

Thus, by general facts about Galois connections:

Corollary 26
(Fin) *$C_{\_}$ is an isomorphism, whose inverse is $\Rightarrow_{\_}$, from $(BCONS_L^{\mathrm{coinf}}, \subseteq)$ onto $(\wp^{\mathrm{coinf}}(Symb_L), \subseteq)$ restricted to minimal sets.*

Assuming (Fin) but without the restriction to co-infinite sets of symbols, there is for every X at least one minimal set generating the same consequence relation as X (Theorem 11), but uniqueness is not guaranteed. With the supplementary assumption that only co-infinite sets are considered, Corollary 26 says that $C_{\Rightarrow_X}$ is the unique minimal set generating the same consequence relation as X.

All the results in this section are made possible by considering only special languages (the rich ones), special consequence relations (the finite or compact ones) and special sets of constants (the co-infinite ones). Rather than making specific assumptions such as these, another way to get results would be to work with a more general definition of $\Rightarrow_{\_}$ that would encapsulate what is necessary to get Lemma 19. This alternative route is explored in the next section.

7 Languages permitting expansions

Richness, or co-infinity, is all about having available symbols *in the language* L. But these symbols are just, as we said, 'stop-over' symbols enabling us to spot logical constants; they play no other role in L. It may seem *ad hoc*, or even unreasonable, to require of an interpreted language that it contain such an unlimited supply of extra symbols. It would be much more reasonable to have a mechanism for *adding* them whenever needed. Instead of L, you go to a suitable *expansion* of L. We shall slightly revise our Bolzano set-up to make this possible, with the aim of eliminating seemingly *ad hoc* assumptions like richness.

This is also a further step towards a Tarskian model-theoretic framework. In such a framework, merely expanding the language is always *conservative* in the sense that the consequence relation for the old language is not affected. As is clear from the previous sections, this may fail drastically in a substitutional setting.[21] We now eliminate this obvious limitation of the classical substitutional framework, while still remaining in a Bolzano style setting.[22]

Why not go directly to the model-theoretic framework? As we said in the Introduction, one reason is that although logical consequence (with respect to a set X of constants) then becomes essentially the familiar notion, it is less clear how to extract constants from consequence relation in a (partly) uninterpreted language. Indeed, the notion of an arbitrary consequence relation in an interpreted language is robust, but much less so if the language is not interpreted (and the constants are not selected in advance). Furthermore, the possibility to add new symbols to a given interpreted language, without disturbing any sentence of the old language, seems like a very mild extension of the Bolzano framework.

There might still be the following worry: If we start with an interpreted language, and then expand this language in various ways, how do we know what the new symbols (and sentences) *mean*? In principle, the answer is: we are free to *stipulate* what they mean, as long as this doesn't 'disturb' the meanings of symbols (and sentences) in L. In fact, we shall see that for the applications in this section, each new symbol we introduce can be taken to be *synonymous* with some L-symbol, in the precise sense that interchanging occurrences of these two symbols never changes the truth values of sentences containing them. Such expansions will be called *expansions with copies*. Then, the extra feature added to the Bolzano framework is merely to allow free introduction of new names for old things.

7.1 Expansions

Recall that, for each language L, $Cat_L = \{C_L : C \in Cat\}$ partitions $Symb_L$.

Definition 27
We say that L' is an *expansion* of L, in symbols $L \leq L'$, iff

$Symb_L \subseteq Symb_{L'}$

For each category $C \in Cat$, $C_L \subseteq C_{L'}$

$Sent_L = \{\varphi \in Sent_{L'} : V_\varphi \subseteq Symb_L\}$

$Tr_L = Tr_{L'} \cap Sent_L$

One easily verifies that

(35) $\leq$ is a partial order (reflexive, antisymmetric, and transitive).

[21] For example, expand the language L_1 by adding a new symbol c such that Rac is false. Then, although $\Rightarrow_{\{a\}} Rab$ holds in L_1, it fails in the expanded language.

[22] The idea is not new; it was proposed, for example, in Bonevac (1985), in the context of first-order logic. There the motivation was to be able to talk about uncountable domains in a countable language with substitutional interpretation of the quantifiers. Here our topic is consequence relations in general and logical constants, but it will be clear from Lemma 42 in Section 8 that in the framework proposed in that section, the consequence relations $\models_{FO_{\mathrm{subst}}}$ and $\models_{FO}$ (Section 3.3) will coincide. Most of Bonevac's paper is about arguing that it is natural to consider expansions in a substitutional setting; we can only agree.

A partially ordered set Z is *directed* iff it is upward closed: if $a, b \in Z$ there is $c \in Z$ such that $a \leq c$ and $b \leq c$. Now, our idea is to start as before with a fixed language L, but also consider a directed family $\mathcal{L}$ of expansions of L. We must then reformulate what have done so far accordingly. First, here is the new notion of Bolzano consequence:

Definition 28
For $\Gamma \cup \{\varphi\} \subseteq Sent_L$ and $X \subseteq Symb_L$: $\Gamma \Rrightarrow_{X,L} \varphi$ iff for every $L' \in \mathcal{L}$ and every replacement ρ in L' (for Γ and φ) which is the identity on X, if $\Gamma[\rho] \subseteq Tr_{L'}$, then $\varphi[\rho] \in Tr_{L'}$.

The family $\mathcal{L}$ is suppressed in this notation, and has to be made clear in context. If $\mathcal{L} = \{L\}$, we have our previous notion of Bolzano consequence: $\Rrightarrow_{X,L} = \Rrightarrow_X$. But in general, $\Rrightarrow_{X,L} \subsetneq \Rrightarrow_X$.

Normally, the sentences we talk about will belong to several languages in $\mathcal{L}$. But since consequence is defined in terms of all expansions (in $\mathcal{L}$) of a given language, this is not a problem. That is, we now have the conservativity property for expansions that fails in the old setting (cf. note 21):

Lemma 29
(Conservativity Lemma) *If* $\Gamma \cup \{\varphi\} \subseteq Sent_L$, $X \subseteq Symb_L$, *and* $L' \in \mathcal{L}$, *then*

$$\Gamma \Rrightarrow_{X,L} \varphi \ \text{ iff } \ \Gamma \Rrightarrow_{X,L'} \varphi$$

or, equivalently,

$$\Rrightarrow_{X,L} \ = \ \Rrightarrow_{X,L'} \restriction Sent_L$$

where the right-hand side is relative to the subfamily $\mathcal{L}' = \{L'' \in \mathcal{L} \colon L' \leq L''\}$.

Proof. If there is a counter-example, via a replacement in some $L'' \geq L'$, to $\Gamma \Rrightarrow_{X,L'} \varphi$, then there is one to $\Gamma \Rrightarrow_{X,L} \varphi$ as well, since $L \leq L''$. Conversely, if there is a counter-example, via a replacement in some $L''' \geq L$, to $\Gamma \Rrightarrow_{X,L} \varphi$, choose, by directedness, L'''' such that $L''' \leq L''''$ and $L' \leq L''''$. Then we have a counter-example in L'''' (with the same replacement) to $\Gamma \Rrightarrow_{X,L'} \varphi$. $\square$

In what follows, when $\mathcal{L}$ is given and $L' \in \mathcal{L}$, we always understand $\Rrightarrow_{X,L'}$ to be relative to the corresponding subfamily generated by L'. Note that each $\Rrightarrow_{X,L'}$ is a consequence relation in L' — i.e. reflexive, transitive, monotone, and truth-preserving — but also, by the Conservativity Lemma, in each expansion of L'. Moreover, relations of this form are base monotone, and (straightforwardly adjusted versions of) the Replacement and Occurrence lemmas hold.

7.2 Useful classes of expansions

Our revised notions of consequence, extraction, etc. (see below) work for any directed class $\mathcal{L}$ of expansions of L. In particular, let

$$exp(L)$$

be the class of *all* expansions of L, and let

$$copies(L)$$

be the class of *expansions with copies* of L, i.e. expansions such that each new symbol is *synonymous*, in the sense indicated above, to some L-symbol. It is straightforward to verify that $(copies(L), \leq)$ is also a directed partial order.

Our toy language L_2 is an expansion of L_1, but not an expansion with copies. To make c a copy of b, both of Rba and Rca must be false, not just Rba as in L_2.

We say that a class $\mathcal{L}$ of expansions of L is *full*, if for all sets $\{a_i : i \in I\} \subseteq Symb_L$ there is an expansion $L' \in \mathcal{L}$ and distinct symbols $b_i \in Symb_{L'} - Symb_L$ of the same category as a_i, for $i \in I$. Clearly, $copies(L)$ (and hence $exp(L)$) is full.

Suppose $\Gamma \Rrightarrow_{X,L} \varphi$, L' is an expansion with copies of L, and Γ', φ' result from Γ, φ by replacing some occurrences of L-symbols with copies in L'. We *cannot* conclude that $\Gamma' \Rrightarrow_{X,L'} \varphi'$, for it may be the case that some but not all occurrences of an L-symbol u have been replaced by a copy $\overline{u}$ (or distinct occurrences by distinct copies), and a replacement of u and $\overline{u}$ by distinct symbols may then yield a counter-example that was not available before. One easily verifies, however, that the following converse holds:

(36) With $L, L', \Gamma, \Gamma', \varphi, \varphi'$ as above: if $\Gamma' \Rrightarrow_{X,L'} \varphi'$, then $\Gamma \Rrightarrow_{X,L} \varphi$.

7.3 General consequence relations

Definition 28 associates with each $X \subseteq Symb_L$ not just one consequence relation, but a *conservative family* of consequence relations, one for each $L' \in \mathcal{L}$. Such families can be seen as instances of a new notion of consequence. As before, $\mathcal{L}$ is a directed family of expansions of a base language L.

Definition 30
A *general consequence relation* (for $\mathcal{L}$) is a family of consequence relations (in the old sense) $\Rrightarrow \; = \{\Rightarrow^{L'}\}_{L' \in \mathcal{L}}$ such that for all $L', L'' \in \mathcal{L}$ with $L' \leq L''$, $\Rightarrow^{L'} \; = \; \Rightarrow^{L''} \restriction Sent_{L'}$.

General consequence relations are partially ordered under a generalized notion of inclusion: define

(37) $\Rrightarrow \sqsubseteq \Rrightarrow'$ iff for all $L' \in \mathcal{L}$, $\Rightarrow^{L'} \subseteq \Rightarrow'^{L'}$

Furthermore, we write, when $\Gamma \cup \{\varphi\} \subseteq Sent_L$,

$$\Gamma \Rrightarrow \varphi$$

instead of $\Gamma \Rightarrow^{L} \varphi$. By conservativity, this is equivalent to $\Gamma \Rightarrow^{L'} \varphi$ holding for all $L' \in \mathcal{L}$ (or for some $L' \in \mathcal{L}$).[23]

[23]The notation is handy, but strictly speaking it means that we are using '$\Rrightarrow$' in two senses: as a family of consequence relations and as a consequence relation. Thus, we employ $\sqsubseteq$ for

General Bolzano consequence relations are of course prime examples of general consequence relations, and we shall write

$$\Rrightarrow_X \;=\; \{\Rrightarrow_{X,L'}\}_{L'\in\mathcal{L}}$$

for the family of consequence relations generated from $X \subseteq Symb_L$ and $\mathcal{L}$ according to Definition 28. General consequence relations of this form satisfy base monotonicity, and the Replacement and Occurrence Lemmas hold.

Next, the notion of minimality is as before: $X \subseteq Symb_L$ is *minimal* iff for each $u \in X$, $\Rrightarrow_X \not\sqsubseteq \Rrightarrow_{X-\{u\}}$. Again it is clear that minimality coincides with being minimal among the sets generating the same general consequence relation.

We say that a general consequence relation $\Rrightarrow \;=\; \{\Rightarrow^{L'}\}_{L'\in\mathcal{L}}$ is *compact* if each $\Rightarrow^{L'}$ is compact. The proof of Theorem 11 in Westerståhl (2010) is easily modified to give:

Theorem 31
For every $X \subseteq Symb_L$, if the general consequence relation $\Rrightarrow_X$ is compact, then X has a subset which is minimal among those generating $\Rrightarrow_X$.

We shall, however, obtain another proof of this theorem in the next subsection.

Finally, we generalize the definition of $C_{\_}$ to general consequence relations of the form $\Rrightarrow = \{\Rightarrow^{L'}\}_{L'\in\mathcal{L}}$. Let $u \in Symb_L$.

Definition 32
$u \in C_{\Rrightarrow}$ iff for some $L' \in \mathcal{L}$, $u \in C_{\Rightarrow^{L'}}$.

Thus, $C_{\Rrightarrow}$ may properly include $C_{\Rightarrow^L}$, since the inference that gets destroyed by replacing u may belong to a proper expansion of L.

We now have the two 'easy' Galois conditions:

(38) a. If $X \subseteq Y$, then $\Rrightarrow_X \sqsubseteq \Rrightarrow_Y$ [base monotonicity]
b. $C_{\Rrightarrow_X} \subseteq X$ [by Replacement as before]

---

the partial order among such families, but

$$\Rrightarrow \subseteq \Rrightarrow'$$

is used as before for the inclusion relation between (ordinary) consequence relations, meaning that if $\Gamma \Rrightarrow \varphi$ then $\Gamma \Rrightarrow' \varphi$; a weaker claim than $\Rrightarrow \sqsubseteq \Rrightarrow'$. Likewise, let us agree to use

$$\Rrightarrow = \Rrightarrow'$$

for equality between the consequence relations (i.e. $\Rightarrow^L = \Rightarrow'^L$), and instead

$$\Rrightarrow \equiv \Rrightarrow'$$

for equality between the families (i.e. for all $L' \in \mathcal{L}$, $\Rightarrow^{L'} = \Rightarrow'^{L'}$).

7.4 The Galois connection liberated

With expansions available, we don't have to worry about sufficiently many symbols being in the base language L. More precisely, Lemma 19 now holds without the restriction to co-infinite sets of symbols or finite sets of sentences.

In the remainder of this section, let $\mathcal{L}$ be any full directed class of expansions of L (Section 7.2).

Lemma 33
If Δ is any set of L-sentences, then for all $X, Y \subseteq Symb_L$, there is an expansion $L' \in \mathcal{L}$ such that in L', X-replacements in Δ factor through Y.

Proof. Let $(Y - X) \cap V_\Delta = \{a_i : i \in I\}$. Since $\mathcal{L}$ is full, some expansion $L' \in \mathcal{L}$ contains for each a_i a distinct symbol b_i outside L of the same category. The rest of the proof is exactly as the proof of Lemma 19. $\square$

As a result, we obtain monotonicity of $C_{\_}$ (Theorem 20) without the previous restrictions.

Theorem 34
$\Rrightarrow_X \sqsubseteq \Rrightarrow_Y$ implies $C_{\Rrightarrow_X} \subseteq C_{\Rrightarrow_Y}$.

Proof. The proof is essentially the same, but we repeat it to indicate the use of expansions. Thus, assume $\Rrightarrow_X \sqsubseteq \Rrightarrow_Y$ and $u \in C_{\Rrightarrow_X}$. We must show that $u \in C_{\Rrightarrow_Y}$. By definition, there is an expansion L' of L in $\mathcal{L}$, and $\Gamma \cup \{\varphi\} \subseteq Sent_{L'}$ and $u' \in Symb_{L'}$ such that $\Gamma \Rrightarrow_{X,L'} \varphi$ but $\Gamma[u/u'] \not\Rrightarrow_{X.L'} \varphi[u/u']$. Thus, there are $L'' \geq L'$ in $\mathcal{L}$ and a replacement ρ in L'' acting outside of X, such that $\Gamma[u/u'][\rho] \subseteq Tr_{L''}$ but $\varphi[u/u'][\rho] \notin Tr_{L''}$.

By Lemma 33 with respect to X, Y and $\Delta = \Gamma \cup \{\phi\} \cup \Gamma[u/u'] \cup \{\phi[u/u']\}$, there is $L''' \geq L''$ in $\mathcal{L}$ such that in L''', X-replacements in Δ factor through Y. So there are σ and τ such that σ acts only on $Y - X$, $\sigma(Y - X) \cap (V_\Delta) = \emptyset$, τ acts outside Y, and $\rho = \tau \circ \sigma$. By Replacement and conservativity,

$$\Gamma[\sigma] \Rrightarrow_{X,L'''} \varphi[\sigma]$$

and thus, by hypothesis,

$$\Gamma[\sigma] \Rrightarrow_{Y,L'''} \varphi[\sigma] \tag{39}$$

Then we can show exactly as before that

$$\Gamma[\sigma][u/\sigma(u')] \not\Rrightarrow_{Y,L'''} \varphi[\sigma][u/\sigma(u')] \tag{40}$$

(39) and (40) entail that $u \in C_{\Rrightarrow_Y}$. $\square$

Similarly, by following the earlier proof, inserting expansions at suitable points, we get a new version of Theorem 21. However, at one crucial step in that proof (the proof of (29)), it is required that $V_{\Gamma \cup \{\varphi\}}$ is finite (not just that it is co-infinite). Therefore, we still need the assumption (Fin), or at least compactness, for this result:

Theorem 35
(FIN) *For every* $X \subseteq Symb_L$, $\Rrightarrow_X \equiv \Rrightarrow_{C_{\Rrightarrow_X}}$.

Interestingly, although compactness plays no role for the monotonicity of $C_{\_}$ in the expansions framework, it cannot be dropped in Theorem 35:

Fact 36
There is a language L *and a full directed class of expansions of* L *with respect to which, for some* $X \subseteq Symb_L$, $\Rrightarrow_X$ *is not compact, and* $\Rrightarrow_X \not\sqsubseteq \Rrightarrow_{C_{\Rrightarrow_X}}$.

Proof. Consider again the language $L_{\mathbb{N}}$ defined in the proof of Fact 12, and let $\mathcal{L} = copies(L_{\mathbb{N}})$. As before, take $X = \{c_0.c_1, \ldots\}$. Now suitable versions of (19) – (21) will hold, in fact with proofs very similar to those for (31) – (33) in the proof of Fact 23; note that the language $L'_{\mathbb{N}}$ considered there was an expansion with copies of $L_{\mathbb{N}}$. We give some indications. First,

(41) If L is an expansion with copies of $L_{\mathbb{N}}$, $\Gamma \subseteq Sent_L$, and infinitely many sentences of the form Qc_i belong to Γ (where Q is a 1-place predicate symbol in L), then $\Gamma \Rrightarrow_{X,L} Inf\, Q$.

The proof is as before, except that a replacement ρ such that all sentences in $\Gamma[\rho]$ are true may now replace Q by a new predicate symbol $\rho(Q)$. But since $\rho(Q)$ must be a copy some $(\neg)P_A$, and since $\rho(Inf)$ must be a copy of *Inf*, it follows in the same way that $\rho(Inf)\rho(Q)$ is true. Next,

(42) If L is an expansion with copies of $L_{\mathbb{N}}$, $Inf\, Q \notin \Gamma$, and if only finitely many sentences of the form Qc_i belong to Γ, then $\Gamma \not\Rrightarrow_{X,L} Inf\, Q$.

(43) If L is an expansion with copies of $L_{\mathbb{N}}$ and $Qd \notin \Gamma$, then $\Gamma \not\Rrightarrow_{X,L} Qd$.

The proof of (43) uses essentially the replacement used for (33) in the proof of Fact 23. Now we can follow the argument for (34) in that proof to establish

(44) $C_{\Rrightarrow_X} = \emptyset$

Then we have (from (41) with $L = L_{\mathbb{N}}$) that $\{P_{\{0\}}c_n : n \in \mathbb{N}\} \Rrightarrow_X Inf\, P_{\{0\}}$. This inference gives a counter-example to compactness as before. But non-compactness also follows from Theorem 35, together with the obvious fact that $\{P_{\{0\}}c_n : n \in \mathbb{N}\} \not\Rrightarrow_{\emptyset} Inf\, P_{\{0\}}$, which establishes that $\Rrightarrow_X \not\sqsubseteq \Rrightarrow_{C_{\Rrightarrow_X}}$. □

The corollaries of Theorems 34 and 35 follow just as in Section 6.3. Let $GBCONS_L$ be the set of general consequence relations of the form $\Rrightarrow_X$ for some $X \subseteq Symb_L$.

Theorem 37
(FIN) $C_{\_}$ *and* $\Rrightarrow_{\_}$ *constitute a Galois connection between* $(GBCONS_L, \sqsubseteq)$ *and* $(\wp(Symb_L), \subseteq)$.

Corollary 38
(FIN) *The image under* $C_{\_}$ *of* $GBCONS_L$ *is the set of minimal sets in* $\wp(Symb_L)$.

Corollary 39
(FIN) *$C_{\text{-}}$ is an isomorphism, whose inverse is $\Rrightarrow_{\text{-}}$, from $(GBCONS_L, \sqsubseteq)$ onto $(\wp(Symb_L), \subseteq)$ restricted to minimal sets.*

Thus, as before, under (FIN) (or compactness) we have shown that $C_{\Rrightarrow_X}$ is the unique minimal set generating the same consequence relation as X. In particular, we have a (new) proof of Theorem 31.

8 From Bolzano to Tarski

Finally, we extend our results in the previous section to cover the more familiar Tarskian semantic notion of logical consequence. This hinges on the fact that substitutional consequence is equivalent to semantic consequence when quantification over expansions is allowed.

8.1 Tarskian interpreted languages

Up to now, our interpreted languages L came equipped only with a set of true sentences. No more was needed to define a substitutional notion of consequence 'à la Bolzano'. For a Tarski style notion of consequence we also need a notion of interpretation for a language and a notion of truth with respect to interpretations. Accordingly, we now introduce *Tarskian interpreted languages*, which come equipped with interpretations, and we assume that a general definition of truth with respect to an interpretation is available for the family of languages under consideration. We shall assume as little as possible regarding the nature of interpretations and the truth relation.

For each (syntactic) category C, let S_C be a corresponding semantic category, intended to be the class of possible semantic values for symbols of category C.

Definition 40
A *Tarskian interpreted language* is a triple $L = \langle Symb_L, Sent_L, I_L \rangle$, where $Symb_L$ and $Sent_L$ are as before, and I_L is an *L-interpretation*, i.e. a function mapping each symbol $u \in Symb_L$ of category C to a semantic value $I(u)$ in S_C. I_L is called the *standard interpretation* of L.

Let $\mathcal{I}_L$ be the class of L-interpretations. We assume that the general truth definition yields, for each Tarskian interpreted language L, a truth relation $\models_L \subseteq \mathcal{I}_L \times Sent_L$. Defining

$$Tr_L = \{\varphi \in Sent_L \colon I_L \models_L \varphi\}$$

we see that Tarskian interpreted languages are special cases of interpreted languages: the case when every symbol has its standard interpretation. Note, however, that in contrast with the more familiar situation in model-theoretic semantics, an interpretation here interprets *all* symbols of the language, not just the 'non-logical' ones.

When I and I' are two interpretations and X is a set of symbols,[24]

$$I =_X I'$$

means that I and I' agree on symbols in X, that is for all $u \in X$, $I(u) = I'(u)$. Our only requirement on the truth definition is that truth should be *local* in the following sense:

(45) If ρ is a replacement such that for all $u \in V_\varphi$, $I(u) = I'(\rho(u))$, then $I \models \varphi$ iff $I' \models \varphi[\rho]$.

Locality means that the question whether a sentence is true or not depends only on the semantic values of its symbols. In particular, if $I =_{V_\varphi} I'$, then $I \models \varphi$ iff $I' \models \varphi$. Arguably, any reasonable truth definition makes truth local — if semantic values do not determine truth values, they are not semantic values. Thus, we do not consider any other component in interpretations over and above semantic values. For example, there is in the present set-up no varying domain of interpretation which could make the truth value of sentences vary even when the semantic values of their symbols remain the same.[25]

8.2 The semantic notion of logical consequence

Tarski's semantic definition of logical consequence as preservation of truth under all possible reinterpretations of non-logical constants can be stated for a Tarskian interpreted language L in the usual way:

Definition 41
φ is a logical consequence of Γ with respect to a set of symbols X,

$$\Gamma \models_{X,L} \varphi\,,$$

iff for all interpretations J such that $J =_X I_L$, if $J \models \Gamma$, then $J \models \varphi$.

Given a Tarskian interpreted language L, substitutional consequence $\Rightarrow_X$ and semantic consequence $\models_{X,L}$ may be compared. As discussed by Tarski himself in Tarski (1936), the substitutional definition makes logical consequence depend on the availability of symbols in L. An inference might be valid just because some semantic values needed to provide a counter-example are not the interpretations of any L-symbols. By contrast, the semantic definition makes all semantic values available by allowing for arbitrary reinterpretation. Thus, $\Gamma \models_{X,L} \varphi$ implies $\Gamma \Rightarrow_X \varphi$, but the converse is not true in general. However,

[24] We drop L as a prefix or subscript when no ambiguity arises.

[25] We take this to be consonant with the original definition of logical consequence in Tarski (1936), which, in contrast to the modern model-theoretic one, was also given for an interpreted language and did not mention varying domains. Actually, the question whether changes in a domain's size were considered by Tarski at the time is a matter of dispute among Tarski scholars; see, for example, Gómez-Torrente (1996). Independently of historical issues, and for the sake of generality, one could think of ways to encode domain variations in the changes of semantic values, but we shall not pursue that here.

the substitutional definition acquires a semantic flavor when expansions come into play as they did in the previous section.

The Tarskian consequence relation $\models_{X,L}$ relative to a set X of constants was defined as usual: there is no need to mention expansions of the interpreted language L since its symbols can be *re*interpreted. But just like other interpreted languages, Tarskian interpreted languages *can* be expanded. The definition is the same as before, except that we require $I_{L'} = I_L \upharpoonright Symb_L$ instead of $Tr_L = Tr_{L'} \cap Sent_L$. Locality guarantees that the former implies the latter. As a consequence, Tarskian expansions are a special kind of expansions. Given a Tarskian language L, we can consider the family

$$exp_T(L)$$

of all its Tarskian expansions. One can easily check that it is a full and directed family. Now substitutional consequence with respect to $exp_T(L)$ becomes equivalent to semantic consequence:

Lemma 42
With respect to $\mathcal{L} = exp_T(L)$, $\Gamma \models_{X,L} \varphi$ *iff* $\Gamma \Rrightarrow_{X,L} \varphi$.

Proof. From left to right: assume $\Gamma \models_{X,L} \varphi$. Let L' be a Tarskian expansion of L and ρ a replacement in L' acting outside X. We need to show that if $I_{L'} \models \Gamma[\rho]$, then $I_{L'} \models \varphi[\rho]$. Define an L'-interpretation J by

$$J(u) = \begin{cases} I_{L'}(\rho(u)) & \text{if } u \in dom(\rho) \\ I_L(u) & \text{otherwise} \end{cases}$$

By definition of J and locality, for any $\psi \in Sent_L$, $J \models \psi$ iff $I_{L'} \models \psi[\rho]$. But $I_L =_X J$, therefore $\Gamma \models_{L,X} \varphi$ implies that if $J \models \Gamma$, then $J \models \varphi$. Hence if $I_{L'} \models \Gamma[\rho]$, then $I_{L'} \models \varphi[\rho]$, as required.

From right to left: assume $\Gamma \Rrightarrow_{X,L} \varphi$. Let J be an interpretation such that $I_L =_X J$. We need to show that if $J \models \Gamma$ then $J \models \varphi$. We define a Tarskian expansion L' by adding a copy u' of each symbol u for which $J(u) \neq I_L(u)$. Copies have the same interpretation as the symbols they are copies of, that is, we set $I_{L'}(u') = J(u)$. Now consider the replacement ρ which maps each such u to u' and is the identity elsewhere. Again, by locality, for any $\psi \in Sent_L$, $J \models \psi$ iff $I_{L'} \models \psi[\rho]$. Therefore, since $I_{L'} \models \Gamma[\rho]$ implies $I_{L'} \models \varphi[\rho]$, $J \models \Gamma$ implies $J \models \varphi$, as required. $\square$

It follows that the general consequence relation $\Rrightarrow_X = \{\Rrightarrow_{X,L'}\}_{L' \in exp_T(L)}$ can also be written $\models_X = \{\models_{X,L'}\}_{L' \in exp_T(L)}$. In this case there is essentially only a notational difference between the consequence relation $\models_{X,L}$ and the family of consequence relations it generates, since each $\models_{X,L'}$ is defined independently of the expansions of L'. Still, since $\models_X$ *can* be seen as a general consequence relation in our sense, and since $exp_T(L)$ is a full family, Theorem 37 applies. Let $TCONS_L$ be the set of general consequence relations of the form $\models_X$ for some $X \subseteq Symb_L$, where L is a Tarskian interpreted language.

Theorem 43
(FIN) *$C_{\_}$ and $\models_{\_}$ constitute a Galois connection between $(TCONS_L, \sqsubseteq)$ and $(\wp(Symb_L), \subseteq)$.*

This happy ending stems from a double virtue of expansions. On the one hand, they allow the Galois connection to hold. On the other hand, they allow semantic consequence to be reduced to substitutional consequence. Even though the problem in both cases amounts to circumventing potential limitations in the richness of the language, expansions do not play exactly the same role in the two cases. To get the equivalence between semantic consequence on the one hand and the substitutional definition of consequence with quantification over expansions on the other, expansions have to be *semantically rich*, they need to provide enough symbols to make all semantic values available. To get the Galois connection, expansions have to be *syntactically rich*, they need to have enough new symbols for the purely syntactic factorization property (Lemma 33) to hold.

9 Further perspectives

9.1 Where we are

The extraction procedure hardwired in the definition of $C_{\_}$ can rightly be taken to satisfy the two adequacy criteria mentioned in the Introduction. First, $C_{\_}$ yields results in accordance with our intuitions when applied to standard examples of logical consequence relations. Second, extraction thus defined does provide an inverse to the process of generating a consequence relation from a set of constants. This claim was made mathematically precise by means of the concept of a Galois connection. In particular, if one allows expansions to play a role in the definition $\Rrightarrow$ of consequence, $C_{\_}$ turns out to constitute a Galois inverse to $\Rrightarrow_X$ (on compact consequence relations defined on a full family of expansions). This was eventually shown to cover the familiar case of compact Tarskian consequence relations. In these settings, the role of $C_{\_}$ on $\Rrightarrow_X$ is to pick out a unique minimal set of constants.

We shall end by noting some potential limitations of the definition of $C_{\_}$ and, reflecting on them, suggest a few leads for further work. Our extraction procedure can be claimed to be both quite liberal and quite severe. We say that u is constant if it occurs essentially in at least one valid inference, in the sense that one can get to an invalid inference by replacing that symbol and nothing else. The phrase 'occurs essentially in *at least one* valid inference' in the definitional clause is responsible for the liberality. Is one inference enough for constancy?[26] The phrase 'by replacing that symbol *and nothing else*' is responsible for the severity. Why should one replace only one thing at a time?

[26] As we noted in Section 1.2, the stronger version (which would read 'in *all* valid inferences') is not easily workable because of the necessary qualification regarding valid inferences whose validity is not due to the purported constant.

9.2 Analytic and logical consequence

As a consequence of the definition of $C_{\_}$ being liberal, $C_{\Rightarrow}$ is bound to declare many more inferences valid than $\Rightarrow$ does, at least for many relations $\Rightarrow$ not of the form $\Rightarrow_X$. The reason is that $\Rightarrow$ might include some meaning postulates for a symbol u, even though it does not treat u as a logical constant. The kind of scenario we have in mind is one where $\Rightarrow$ *partly* fixes the interpretation of a symbol u by declaring valid some inferences essentially involving u, but cannot be construed as being of the form $\Rightarrow_X$ for some X with $u \in X$. In such a scenario, u will belong to $C_{\Rightarrow}$, so that $\Rightarrow_{C_{\Rightarrow}}$, contrary to $\Rightarrow$, relies on keeping the denotation of u *completely* fixed.

The fact that $\Rightarrow_{C_{\Rightarrow}}$ properly extends $\Rightarrow$ may not be a problem *per se*. The problem is that $C_{\_}$ cannot be used to tell the difference between logical inferences and merely analytic inferences. One might have hoped that C would select logical constants, in a way such that the further application of $\Rightarrow_{\_}$ would have isolated a core of purely logical inferences. Thus, $\Rightarrow_{C_{\Rightarrow}}$ would have been a subset of $\Rightarrow$, the subset of its purely logical inferences, whereas inferences in $\Rightarrow - \Rightarrow_{C_{\Rightarrow}}$ would have been the analytic ones.

Failure of $C_{\_}$ to do this job is particularly unfortunate with respect to the intended application to (logically uneducated) speakers of a language, if identification of the logical apparatus of the language is the goal. Speakers can hardly be assumed to be able by themselves to tell the difference between logical and analytic inferences. The only notion $\Rightarrow$ of valid inference on which a linguist can rely is bound to include both logical and analytic inferences. In that circumstance, $C_{\Rightarrow}$ will overgenerate by encompassing every basic expression to which a meaning postulate is attached.

A solution to this problem would consist in finding ways to further filter the results given by $C_{\_}$. A difference between analytic inferences and logical inferences is that the latter but not the former are schematic. An analytic inference is specific to the expressions involved, whereas a logical inference is generally valid. One idea would be to retain only those constants in $C_{\_}$ whose valid inferences are schematic *with respect to* $\Rightarrow$.[27]

9.3 Non-uniform consequence

Let us turn to $C_{\_}$ being too severe. In principle, it seems that nothing precludes the role played by a constant u to show up only in connection with other substitutions. In that case, $C_{\_}$ would fail to select u. By contrast, one might consider a different extraction procedure, say $C^*_{\_}$. As before, $u \in C^*_{\Rightarrow}$ would require finding a valid inference $\Gamma \Rightarrow \varphi$ in which u occurs. But now the invalid inference which is to witness u's essential involvement in that validity could be obtained by means of a replacement ρ which moves u (as before) but possibly other symbols as well. However, this cannot be the whole story. It could not yet capture the fact that u was essential to the validity of $\Gamma \Rightarrow \varphi$, since, after

[27] A definition of schematicity has not been given, but it could rely on $C_{\Rightarrow}$ to state the necessary restriction on the range of replacements.

all, putting some other symbol in place of u could be totally contingent to the destruction of the inference. We need to require that substituting u was indeed necessary for ρ to do so. This leads to the following definition: $u \in C^*_{\Rightarrow}$ iff there is an inference $\Gamma \Rightarrow \varphi$ and a replacement ρ such that $\Gamma[\rho] \not\Rightarrow \varphi[\rho]$ but $\Gamma[\rho_{-u}] \Rightarrow \varphi[\rho_{-u}]$, where ρ_{-u} is the replacement which differs from ρ at most on u and maps u to itself.

$C^*_{\_}$ is indeed less severe than $C_{\_}$. It is easy to check that, for any $\Rightarrow$, $C_{\Rightarrow} \subseteq C^*_{\Rightarrow}$.[28] The converse is not true in general, as witnessed by our language L_1. We had $a \notin C_{\Rightarrow_{\{a\}}}$ but we get $a \in C^*_{\Rightarrow_{\{a\}}}$. To see this, recall that $\Rightarrow_{\{a\}} Rab$ (and $\Rightarrow_{\{a\}} Raa$) but $\not\Rightarrow_{\{a\}} Rba$. Let ρ swap a and b. We get $\Rightarrow_{\{a\}} Rab$, $\not\Rightarrow_{\{a\}} Rab[\rho]$ and $\Rightarrow_{\{a\}} Rab[\rho_{\{-a\}}]$. Not uninterestingly, this suggests that C^* solves some of the problems on account of which we had to introduce rich languages or full expansions. By way of ρ, no stop-over is needed, so that $a \in C^*_{\Rightarrow_{\{a\}}}$ not only in the context of L_2 (where c is available) but already in L_1 (where no symbol different from a and b is available).

We shall not engage here in a thorough examination of the properties of $C^*_{\_}$. Despite what was pointed out in the previous paragraph, $C^*_{\_}$ does not yield a straightforward Galois correspondence for $\Rightarrow$ or $\models_X$. However, let us mention that it can be shown that $C^*_{\_}$ does yield a straightforward Galois connection for another notion of logical consequence. The notion we have in mind is stronger than the standard one in that it allows for *non-uniform* replacements of non-logical constants. Accordingly, a classical tautology such as $p \vee \neg p$ ceases to be valid, since $p \vee \neg q$ is not valid. This stronger notion of logical consequence has recently received a lot of interest from linguists who are looking for a connection between logicality and grammaticality (not all validities or contradictions are ungrammatical, but validity or contradiction seems to play a role in some sentences being ungrammatical).[29] It is a rather pleasant surprise that this notion independently appears in connection with the extraction problem and the contrast between $C_{\_}$ and $C^*_{\_}$.

References

Abrusán, M. (2008). Presuppositional and negative islands: a semantic account. *Natural Language Semantics*, to appear.

Aczel, P. (1990). Replacement systems and the axiomatization of situation theory. In R. Cooper, K. Mukai, and J. Perry, editors, *Situation Theory and its Applications, Vol 1*, pages 3–33. CLSI Publications, Stanford.

Bar-Hillel, Y. (1950). Bolzano's definition of analytic propositions. *Theoria*, **16:2**, 91–117.

[28] As a consequence, $C^*_{\_}$ cannot help with the conceptual difficulties surrounding the difference between logical and analytic inferences.

[29] The idea was introduced by Gajewski under the name of *L-analyticity* in connection with 'there' sentences and exceptives (Gajewski (2002)). It has then been taken up to help explain various other phenomena, including measurement scales (Fox and Hackl (2006)) and presuppositional or negative islands (Abrusán (2008)).

Bolzano, B. (1837). *Theory of Science*. ed. Jan Berg, D. Reidel, Dordrecht, 1973.

Bonevac, D. (1985). Quantity and quantification. *Noûs*, **19**, 229–247.

Bonnay, D. (2008). Logicality and invariance. *Bulletin of Symbolic Logic*, **14:1**, 29–68.

Bonnay, D. and Westerståhl, D. (2010). Logical consequence inside out. In M. Aloni and K. Schulz, editors, *Amsterdam Colloquium 2009*, pages 193–202. LNAI 6042, Springer, Heidelberg.

Carnap, R. (1937). *The Logical Syntax of Language*. Kegan, Paul, Trench Trubner & Cie, London. Rev. ed. translation of *Logische Syntax der Sprache*, Wien, Springer, 1934.

Dunn, M. and Belnap, N. (1968). The substitution interpretation of the quantifiers. *Noûs*, **4**, 177–185.

Feferman, S. (2010). Set-theoretical invariance criteria for logicality. *Notre Dame Journal of Formal Logic*, to appear.

Fox, D. and Hackl, M. (2006). The universal density of measurement. *Linguistics and Philosophy*, **59:5**, 537–586.

Gajewski, J. (2002). L-analyticity and natural language. Manuscript.

Gómez-Torrente, M. (1996). Tarski on logical consequence. *Notre Dame Journal of Formal Logic*, **37:1**, 125–151.

MacFarlane, J. (2009). Logical constants. In E. N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Fall 2009 edition.

Peters, S. and Westerståhl, D. (2006). *Quantifiers in Language and Logic*. Oxford University Press, Oxford.

Tarski, A. (1936). On the concept of logical consequence. In *Logic, Semantics, Metamathematics*, pages 409–420. Hackett Publishing, Indianapolis, 1983.

Tarski, A. (1986). What are logical notions? *History and Philosophy of Logic*, **7**, 145–154.

van Benthem, J. (2003). Is there still logic in Bolzano's key? In E. Morscher, editor, *Bernard Bolzanos Leistungen in Logik, Mathematik und Physik Bd. 16*, pages 11–34. Academia Verlag, Sankt Augustin.

Westerståhl, D. (2010). From constants to consequence, and back. *Synthese*, forthcoming in a special issue on the Philosophy of Logical Consequence and Inference.