THE REVIEW OF SYMBOLIC LOGIC
Volume 1, Number 3, October 2008

BELIEF REVISION IN NON-CLASSICAL LOGICS

DOV GABBAY and ODINALDO RODRIGUES
Department of Computer Science, King's College London

and

ALESSANDRA RUSSO
Department of Computing, Imperial College

Abstract. In this article, we propose a belief revision approach for families of (non-classical) logics whose semantics are first-order axiomatisable. Given any such (non-classical) logic $\mathscr{L}$, the approach enables the definition of belief revision operators for $\mathscr{L}$, in terms of a belief revision operation satisfying the postulates for revision theory proposed by Alchourrón, Gärdenfors and Makinson (AGM revision, Alchourrón *et al.* (1985)). The approach is illustrated by considering the modal logic K, Belnap's four-valued logic, and Łukasiewicz's many-valued logic. In addition, we present a general methodology to translate algebraic logics into classical logic. For the examples provided, we analyse in what circumstances the properties of the AGM revision are preserved and discuss the advantages of the approach from both theoretical and practical viewpoints.

1. Introduction. This article presents a method for revising theories of non-classical logics. The general idea is to define a belief revision operator for these logics in terms of a standard belief revision operator for classical logic. Typically, given a classical logic theory Δ and an input formula ψ, a revision process $*_a$ gives a new theory $\Gamma = \Delta *_a \psi$, corresponding to the revision of Δ by ψ and satisfying some desirable properties such as the postulates for revision theory proposed by Alchourrón, Gärdenfors and Makinson—henceforth called an *AGM revision* (Alchourrón *et al.*, 1985; Gardenfors, 1988). Our approach shows that this machinery can be *exported* to other (non-classical) logics whenever the semantics of these logics can be axiomatised in first-order logic.

This is done by translating the mechanics of a given object logic L into classical logic, performing the revision process in classical logic and then translating the results back into the object logic. Let τ denote a translation function from L into classical logic, $\mathscr{A}_L$ a sound and complete axiomatisation of the semantic features of L in classical logic, and Acc a characterisation of *acceptable* L-theories in first-order logic. We define a revision operator $*_L$ in the logic L as follows:

$$\Delta *_L \psi = \{\beta \mid \Delta^\tau *_a (\psi^\tau \wedge \mathscr{A}_L \wedge Acc) \vdash \beta^\tau\}, \tag{1}$$

where Δ^τ and ψ^τ are the classical logic translations of Δ and ψ, respectively. The revision of Δ by ψ is therefore defined in terms of the classical revision of Δ^τ by ψ^τ. The inclusion of the axiomatisation $\mathscr{A}_L$ allows the semantic properties of the object logic L to be preserved during the revision process, so that the revised theory can be mapped back to L. This is why we have defined the revision of Δ^τ with respect to $\psi^\tau \wedge \mathscr{A}_L$.

Notice that we have also included the formula (Acc) in the revising part of Equation (1). For non-classical logics that are extensions of classical logic, for example, modal logics,

Received: 18 December, 2007

© 2008 Association for Symbolic Logic
doi:10.1017/S1755020308080246

https://doi.org/10.1017/S1755020308080246 Published online by Cambridge University Press

the notion of inconsistency is identical to the notion of inconsistency in the target (classical) logic. If a given formula ψ is inconsistent with a given theory Δ, then ψ^τ would also be inconsistent (in classical logic) with Δ^τ. Hence, the classical revision would be able to revise the (translated) theory and provide a revised theory that is also consistent in the object logic. For these types of non-classical logics, the revision operator $*_L$ could be defined more simply as:

$$\Delta *_L \psi = \{\beta \mid \Delta^\tau *_a (\psi^\tau \wedge \mathscr{A}_L) \vdash \beta^\tau\}.$$

This is the approach adopted and discussed in Gabbay *et al.* (2000). However, even in these cases, some domain-dependent notions of acceptability could still be provided to further refine the revision process, as it is done, for instance, with integrity constraints in database updates. Katsuno & Mendelzon (1991) defined the update of a knowledge base Δ by a formula ψ verifying integrity constraints IC as $\Delta \circ (\psi \wedge IC)$. The difference to our approach is that IC is represented in the same language as that of the object logic, that is, the language of the knowledge base Δ, whereas in our case, the acceptability condition Acc is represented in the target language.[1]

Inconsistency in classical logic is associated with the non-existence of valuations that make formulae true—the non-existence of *models*. In the case of logics that are not extensions of classical logic, for example, Belnap's four-valued logic, *some* notion of 'consistency' has to be specifically defined and will usually differ from that of classical logic. 'Inconsistent' object theories may not necessarily yield inconsistent classical logic theories. For instance, the formula $p \wedge \neg p$, which takes both the truth-values *true* and *false* in Belnap's four-valued logic, will have valuations with only the truth-value *true* in the translation to classical logic. These valuations simply preserve the semantic meaning of $p \wedge \neg p$ in Belnap's logic in classical logic. In other words, they constrain the relationship between the truth-values of p, $\neg p$, and $p \wedge \neg p$, which is entirely determined by the relationship between the values of the atomic formula p and the derived truth-values of $\neg p$ and $p \wedge \neg p$ according to the intended meaning of the connective $\wedge$ in the object logic. If p has the value *true* in a (Belnap) valuation, then $\neg p$ will have the value *false* in that valuation and $p \wedge \neg p$ the value *both*. A model of the translation of these statements into classical logic is simply a valuation in classical logic that preserves these relationships.

It is hence without surprise that if a translation mechanism τ into classical logic is sound and complete, then p^τ will be classically consistent with $(\neg p)^\tau$, and consequently, the revision of p by $\neg p$ done via classical logic will be simply $\{p \wedge \neg p\}$. This might not be what one expects in the object logic. In these circumstances, the revision would be equivalent to an expansion, similarly to what is defined for paraconsistent logics (Restall & Slaney, 1995). An alternative is to embed the revision mechanism in the logic itself, an example of which can be found in Martins & Shapiro (1998) and Shapiro (1992) for the system **R**.

Definition (1) provides the general pattern for the revision by translation methodology. The particular details of the actual translation and notion of acceptability will vary according to the specific object logic.

What we propose in this article is a different approach to revision for non-classical logics (including paraconsistent logics). Our view is that a revision process for such logics

[1] We could of course revise by integrity constraints IC represented in the *object* language by revising by $\psi^\tau \wedge \mathscr{A}_L \wedge IC^\tau$ instead.

https://doi.org/10.1017/S1755020308080246 Published online by Cambridge University Press

would benefit from a shift towards a notion of *acceptable belief sets*, by imposing specific restrictions on what an acceptable theory should be in the object logic L, and therefore on what is the effect of revising a theory with respect to *a* notion of 'acceptability'. The only restriction is that that notion can be expressed in classical logic. We assume that such notion can be provided either by the specific domain of applications or by logical properties, with the aim of investigating the relationship between these properties and the underlying classical AGM postulates. In the case of Belnap's logic, we provide a basic notion used simply to illustrate our approach, and we leave more specific considerations for future work. For a discussion on acceptability in paraconsistent logic, the reader is referred to Tanaka (1997), Priest (2001), and Restall & Slaney (1995).

An earlier investigation of this topic was given in Gabbay *et al.* (2000). This article significantly extends that investigation by considering the case of Łukasiewicz's many-valued logic, by providing a general translation method of algebraic logics into classical logic and considering revision in this class of logics, and by providing a generalisation of Belnap's notion of entailment to infinite theories.

The article is structured as follows. In Section 2, we provide a brief introduction to belief revision in classical logic according to the AGM postulates. In Section 3, we define and motivate our general approach for belief revision in non-classical logics, which is then applied to the modal logic K (Section 4) and Łukasiewicz's logic (Section 5). In Section 6, we present the methodology for algebraic logics in general. We then consider Belnap's four-valued logic, which does not have a notion of consistency and show how our revision approach can be applied in this case (Section 7). Finally, we present some conclusions in Section 8.

2. Belief revision in classical logic. The term *belief revision* is generally used to describe a process for changing a given logical theory to accommodate new (possibly contradictory) information, in such a way that the process will satisfy some desirable properties. One important feature is that it is intrinsically a non-monotonic form of reasoning.

Standard belief revision operations for classical logics assume the given theories to be closed under the consequence relation of the underlying logic, that is, to be *belief sets*. Variants of these operations, called *base revisions*, have also been considered in the literature, where theories are just finite sets of beliefs, that is, *belief bases*. However, in this article, we concentrate only on the standard approach to belief revision. If Cn is the consequence relation of a given logic and K a belief set, then we assumed that $K = \mathrm{Cn}(K)$. Similarly, if for a belief φ and a belief set K, $\varphi \in K$, we say that φ is *accepted* in K.

The general task of the revision process can be seen as determining what is rational to believe after a belief change occurs. The change could be, for instance, the acceptance of new (possibly contradicting) belief.

According to the AGM theory (Gärdenfors, 1988), belief changes can be of three main types:

- *Expansion*: When new information is consistent with the current belief set. In this case, the revised theory is given by the transitive closure over the consequence relation, of the union of the previous belief set with the new sentence.
- *Contraction*: When some information is forced to be retracted from the belief set. Since the belief set is closed under the consequence relation, to retract a belief φ, it is also necessary to remove other beliefs that imply φ.

https://doi.org/10.1017/S1755020308080246 Published online by Cambridge University Press

- *Revision*: Which is the acceptance of new (consistent) information contradicting the current belief set. Revision is in this case the process of restoring consistency over the new belief set whenever the new information is not itself contradictory.

The most interesting cases are contractions and revisions. These have sometimes been defined in terms of each other, for instance, *Levi Identity* has defined the revision process in terms of a primitive operation of contractions and the *Harper Identity* has defined contraction in terms of a primitive operation of revision (Gärdenfors, 1988, pp. 69–70). For the rest of this article, we concentrate mainly on the revision operation.

The AGM postulates have often been considered as standard criteria for evaluating revision operations and provide, in a sense, a *minimal* set of conditions to be satisfied by such operations. We use them to investigate the properties of the operator obtained in our approach. The postulates, as given in Gärdenfors (1988), are described below. In the following presentation, K denotes a belief set, A and B formulae, and $K_\perp$ the inconsistent belief set (the set of all well-formed formulae in the language).

2.1. AGM postulates for belief revision.

(K*1) $K *_a A$ is a belief set.
(K*2) $A \in K *_a A$.
(K*3) $K *_a A \subseteq \mathrm{Cn}(K \cup \{A\})$.
(K*4) If $\neg A \notin K$, then $\mathrm{Cn}(K \cup \{A\}) \subseteq K *_a A$.
(K*5) $K *_a A = K_\perp$ only if A is contradictory.
(K*6) If $A \equiv B$, then $K *_a A \equiv K *_a B$.
(K*7) $K *_a (A \wedge B) \subseteq \mathrm{Cn}((K *_a A) \cup \{B\})$.
(K*8) If $\neg B \notin K *_a A$, then
$\mathrm{Cn}(K *_a A \cup \{B\}) \subseteq K *_a (A \wedge B)$.

Postulate (K*1) requires that the result of the revision operation also be a belief set. (K*2), known as the *success postulate*, requires that the new information is accepted in the revised belief set. Postulates (K*3) and (K*4) define in a sense the 'lower' and the 'upper' bounds of the revision operation, respectively. In Gabbay *et al.* (2000), we have argued that together they amount to the following principle:

($\mathrm{K}^*_{3,4}$) If A is consistent with K, then
$K *_a A = \mathrm{Cn}(K \cup \{A\})$.

(K*5), instead, only requires K to be consistent whenever A is consistent. When we check the AGM postulates for logics other than classical, we may have a different notion of consistency and so we are free to interpret what we want the revision to do in the case of inconsistency according to what is reasonable in the object (non-classical) logic. AGM for classical logic gives us no clue beyond (K*5) as to what to require when $(K *_a A) \cup \{B\}$ is inconsistent. (K*6) specifies that the revision process should be independent of the syntactic form of the sentences involved. Finally, Postulates (K*7) and (K*8) deal with the relationship of inputing (A, B) as a sequence (first revising by A and then expanding by B), compared with revising by $\{A, B\}$ at the same time (i.e., revising by $A \wedge B$). It is well-known that AGM does not say enough about sequences of revisions and their properties. These postulates are the bare minimum (see, for instance, Darwiche & Pearl, 1996, 1997). In fact, (K*7) and (K*8) do not tell us anything new (beyond what we can deduce from earlier postulates), except in the case where B is consistent with $K *_a A$. In Gabbay *et al.*

https://doi.org/10.1017/S1755020308080246 Published online by Cambridge University Press

(2000), we have argued that (K*7) and (K*8) together are equivalent to the following postulate:

$(\mathrm{K}^*_{7,8})$ $\mathrm{Cn}((K *_a A) \cup \{B\}) = K *_a (A \wedge B)$, when B is consistent with $K *_a A$.

For non-classical logics, we are committed only to $(\mathrm{K}^*_{7,8})$. Other cases involving inconsistency can have properties dictated by the object logic requirements.

3. Belief revision in non-classical logics. In this section, we describe our approach for defining belief revision operators for non-classical logics. This is based on the following components:

(i) A sound and complete classical logic axiomatisation of the semantics of the object logic L.
(ii) A domain-dependent notion of acceptability for theories of L.
(iii) A classical AGM belief revision operation.

In general, different translation mechanisms can be defined from a given object logic to classical logic, depending on the underlying semantic structures used by the object logic. In the case of non-classical logics that are extensions of classical logic, the translation mechanism essentially maps the notion of satisfiability of atomic formulae to associated first-order predicates with one or more (additional) arguments to encapsulate elements of the underlying structure of the semantics. For instance, in the case of propositional modal logic, the semantic notion $w \Vdash p$, for a possible world w and atomic formula p can be associated with a first-order formula $P(w)$, and analogously, $w \nVdash p$, with the formula $\neg P(w)$. Appropriate axioms may have to be defined to express the correct behaviour of the connectives (and indeed semantics) of the object logic, for example, special conditions of the accessibility relation (Ohlbach, 1991).

For non-classical logics, whose semantics are significantly different from those of classical logic, the translation mechanism has to mirror the structures on which the semantics as well as the notion of entailment are based. In the case of Belnap's four-valued logic, the semantic notion of entailment is based on a lattice with four elements, called the *logical lattice* (**L4**). The translation mechanism needs to reflect the structure of the lattice and express (declaratively) the different truth-value assignments that a formula can assume under given valuations. An axiomatisation expressing the behaviour of the connectives needs also to be defined. Independently of the translation mechanism adopted, such axiomatisation has to be *sound* and *complete* with respect to the semantic notion of entailment of the object logic L. This ensures that given a translation function τ and a classical logic axiomatisation $\mathscr{A}_L$, for any given theory Δ and formula α in L,

$$\Delta \vdash_L \alpha \text{ iff } \mathscr{A}_L \cup \Delta^\tau \vdash \alpha^\tau,$$

where Δ^τ and α^τ are the classical translations of Δ and α, respectively.[2] Throughout the article, we refer to $\mathscr{A}_L$ as a sound and complete axiomatisation of L in classical logic. Our method is thus applicable only to logics whose semantics are first-order logic axiomatisable.

[2] We use the symbol $\vdash$ without subscript to denote entailment in classical logic and with subscript to denote entailment on a object logic.

https://doi.org/10.1017/S1755020308080246 Published online by Cambridge University Press

The second component of our approach is a domain-dependent notion of *acceptability* for a theory in L. Given a revision operator $*_a$ for classical logic, we could define a revision operator $*_L$ for the logic L in the following way:

$$\Delta *_L \psi = \{\beta \mid \Delta^\tau *_a (\psi^\tau \wedge \mathscr{A}_L) \vdash \beta^\tau\}. \tag{2}$$

However, the notion of consistency in the object logic may differ from that of classical logic. The AGM revision relies somehow on the notion of (classical) consistency. This can be easily understood by analysing Postulate ($K^*_{3,4}$). Revision is only triggered when the new information is inconsistent with the current belief set. In this sense, revision in the translation can only be performed if the translated theory turns out to be *classically* inconsistent. The direct application of our method to logics such as Belnap's four-valued logic, whose notion of inconsistency of a theory is not associated with the non-existence of models of the classical logic translation of that theory, would therefore yield simply to an expansion since the translation of any Belnap theory is always classically consistent. This is a consequence of the fact that AGM belief revision is too dependent on the mechanics of classical logic itself.

Our view is that belief revision processes would benefit from a shift towards the more general notion of *acceptable* belief sets. In the particular case of classical logic, it makes sense to require that an acceptable belief set be *at least* consistent since inconsistency trivialises the notion of consequence in classical logic. However, there is no reason why more specific restrictions should not be imposed on what an acceptable belief set is in classical logic, as it is done, for instance, with integrity constraints in database updates. This paves the way for the definition of *relative* revision procedures, where the process is guided by a fixed theory representing a particular notion of acceptable belief sets. For the purposes of this article, our relative revision is defined in terms of a notion of acceptability explicitly included in the revision process. Our definition of belief revision for a non-classical logic L is therefore as follows.

DEFINITION 3.1 (**Belief revision in L**). *Let $*_a$ be a revision operator for classical logic and Acc a formula expressing a (domain dependent) notion of acceptability of L-theories. The revision of Δ by ψ, $\Delta *_L \psi$, is defined as:*

$$\Delta *_L \psi = \{\beta \mid \Delta^\tau *_a (\psi^\tau \wedge \mathscr{A}_L \wedge Acc) \vdash \beta^\tau\}.$$

Hence, a theory Δ can be revised by a formula ψ, by instead revising classically Δ^τ by ψ^τ. The semantic properties of the object logic L and the notion of acceptability are preserved by performing the classical revision process also with respect to the axiomatisation $\mathscr{A}_L$ and the formula Acc, to guarantee that the resulting revised theory is an acceptable theory in the object logic L.

4. Revising the modal logic K. In this section, we refer to a standard presentation of the modal logic K, for example, Chellas (1980). In particular, we make the use of the notion of entailment in which $\Delta \vdash_k \beta$ if and only if in every Kripke model with actual world w_0, we have $w_0 \Vdash_k \Delta$ implies $w_0 \Vdash_k \beta$.

4.1. Translation of the modal logic K into classical logic. We perform the translation of the model logic K into classical logic via a binary predicate R in classical logic to represent the accessibility relation and unary predicates $P_1, P_2, P_3, \ldots$, for each propositional symbol p_i of K.

We use the subscript k whenever we wish to emphasise that we mean an operation (relation) in k and differentiate it from its classical logic counterpart (which is not subscripted).

https://doi.org/10.1017/S1755020308080246 Published online by Cambridge University Press

The idea is to encode the information of satisfiability of modal formulae by worlds into the variable of each unary predicate. In general, for a given world w and formula β, the translation method can be stated as follows, where $\beta^\tau(w)$ represents $w \Vdash_k \beta$.

$$\begin{aligned} p_i^\tau(w) &= P_i(w) \\ (\neg\beta)^\tau(w) &= \neg(\beta^\tau(w)) \\ (\beta \wedge \gamma)^\tau(w) &= \beta^\tau(w) \wedge \gamma^\tau(w) \\ (\beta \rightarrow \gamma)^\tau(w) &= \beta^\tau(w) \rightarrow \gamma^\tau(w) \\ (\Box\beta)^\tau(w) &= \forall y(wRy \rightarrow \beta^\tau(y)). \end{aligned}$$

Finally, for a modal theory Δ, we define:

$$\Delta^\tau(w) = \{\beta^\tau(w) \mid \beta \in \Delta\}.$$

We then have the following correspondence.

THEOREM 4.1 (**Correspondence**).

$$\Delta \vdash_k \beta \ \textit{iff} \ \mathscr{A}_L \cup \Delta^\tau(w_0) \vdash \beta^\tau(w_0). \tag{3}$$

In the case of the logic K, the theory $\mathscr{A}_L$ is empty (i.e., Truth).[3] If Δ is finite, we can let $\delta = \bigwedge \Delta$, and we have:

$$\delta \vdash_k \beta \ \text{iff} \ \vdash \forall x(\delta^\tau(x) \rightarrow \beta^\tau(x)).$$

4.2. Revision in the modal logic K. We can define a revision operator $*_k$ for K, as outlined before (we omit the reference to the actual world w_0 in the rest of this section).

DEFINITION 4.2 (**Belief revision in K**).

$$\Delta *_k \psi = \{\alpha \mid \Delta^\tau *_a (\psi^\tau \wedge \mathscr{A}_L) \vdash \alpha^\tau\}.$$

We can now speak more specifically of properties of $*_k$.

4.2.1. Properties of the operator $*_k$.

1. $\Delta *_k \psi$ is closed under $\vdash$.
   This can be easily shown.
2. $\Delta *_k \psi \vdash_k \psi$.
   By (K*2), $\psi^\tau \wedge \mathscr{A}_L \in \Delta^\tau *_a (\psi^\tau \wedge \mathscr{A}_L)$. Since $\Delta^\tau *_a (\psi^\tau \wedge \mathscr{A}_L)$ is closed under $\vdash$, $\Delta^\tau *_a (\psi^\tau \wedge \mathscr{A}_L) \vdash \psi^\tau$ and hence $\psi \in \Delta *_k \psi$, by (K*1), $\Delta *_k \psi \vdash_k \psi$.
3. If ψ is (modally) consistent with Δ, then $\Delta *_k \psi = \mathrm{Cn}_k(\Delta \cup \{\psi\})$.
   We first show that if ψ is modally consistent with Δ, then $\Delta^\tau(k)$ is classically consistent with $\psi^\tau(k) \wedge \mathscr{A}_L$. This holds because $\Delta \cup \{\psi\}$ has a Kripke model, which will give rise to a classical model of the translation. Therefore, $\Theta = \Delta^\tau(k) *_a (\psi^\tau(k) \wedge \mathscr{A}_L)$ is the classical provability closure of $\Delta^\tau(k) \cup \{\psi^\tau(k) \wedge \mathscr{A}_L\}$.
   We now have to show that if $\alpha^\tau(k) \in \Theta$, then $\Delta *_k \psi \vdash \alpha$.

LEMMA 4.3. *Let Δ be a closed theory. Let Δ^τ be its translation and let* $\mathrm{Cn}(\Delta^\tau)$ *be its $\mathscr{A}_L$-closure in classical logic. Let β be such that $\beta^\tau \in \mathrm{Cn}(\Delta^\tau)$. It follows that $\beta \in \Delta$.*

[3] The logic K imposes no properties on R. Had we been translating $S4$, we would have $\mathscr{A}_L = \{\forall x(xRx) \wedge \forall x\forall y\forall z(xRy \wedge yRz \rightarrow xRz)\}$. Our notion also allows for non-normal logics, for example, if w_0 is the actual world, we can allow reflexivity in w_0 by setting $\mathscr{A}_L = \{w_0 R w_0\}$.

https://doi.org/10.1017/S1755020308080246 Published online by Cambridge University Press

Proof. If $\beta \notin \Delta$, then there exists a Kripke model of $\Delta \cup \{\neg\beta\}$. This gives a classical model of $\Delta^\tau \cup \{\mathscr{A}_L\} \cup \{\neg\beta^\tau\}$, and so $\beta^\tau \notin \mathrm{Cn}(\Delta^\tau)$. □

LEMMA 4.4. *Let Δ^τ be a closed classical theory such that $\Delta^\tau \vdash \mathscr{A}_L$ and let $\Delta = \{\beta \mid \beta^\tau \in \Delta^\tau\}$. Then if $\Delta \vdash \alpha$, then $\alpha^\tau \in \Delta^\tau$.*

Proof. If $\alpha^\tau \notin \Delta^\tau$, there exists a model of $\Delta^\tau \cup \{\neg\alpha^\tau\}$. This can be viewed as a Kripke model of $\Delta \cup \{\neg\alpha\}$. □

4. $\Delta *_k \psi$ is modally inconsistent, only if ψ is modally contradictory.
   If $\Delta *_k \psi$ is inconsistent, then so is $\Delta^\tau *_a (\psi^\tau \wedge \mathscr{A}_L)$ since $\Delta^\tau *_a (\psi^\tau \wedge \mathscr{A}_L)$ is closed under $\vdash$. By (K*5), $\psi^\tau \wedge \mathscr{A}_L$ is contradictory. But by Theorem 4.1, $\psi^\tau \wedge \mathscr{A}_L \vdash \bot$ if and only if $\psi \vdash_k \bot$.
5. If $\psi \equiv_k \varphi$, then $\Delta *_k \psi \equiv \Delta *_k \varphi$.
   By Theorem 4.1 and the fact that $\psi \vdash_k \varphi$ and $\varphi \vdash_k \psi$, it follows that $\mathscr{A}_L \cup \psi^\tau \equiv \mathscr{A}_L \cup \varphi^\tau$. Therefore, by (K*6), $\Delta *_a (\mathscr{A}_L \cup \psi^\tau) \equiv \Delta *_a (\mathscr{A}_L \cup \varphi^\tau)$ and hence $\Delta *_k \psi \equiv \Delta *_k \varphi$.
6. $\Delta *_k (\psi \wedge \varphi) \subseteq \mathrm{Cn}_k((\Delta *_k \psi) \cup \{\varphi\})$.
   Suppose that $\Delta *_k (\psi \wedge \varphi) \vdash_k \alpha$, for some α. By the definition of $*_k$, $\Delta^\tau *_a (\psi^\tau \wedge \varphi^\tau \wedge \mathscr{A}_L) \vdash \alpha^\tau$. By (K*7), it follows that $\mathrm{Cn}(\Delta^\tau *_a (\psi^\tau \wedge \mathscr{A}_L) \cup \{\varphi^\tau\} \vdash \alpha^\tau$. Notice that for every $\gamma^\tau \in \Delta^\tau *_a (\psi^\tau \wedge \mathscr{A}_L)$, there is a corresponding γ in $\Delta *_k \psi$ (by the definition of $*_k$) and similarly for φ^τ. By Theorem 4.1, $\mathrm{Cn}_k((\Delta *_k \psi) \cup \{\varphi\}) \vdash \alpha$.
7. If φ is modally consistent with $\Delta *_k \psi$, then $\mathrm{Cn}((\Delta *_k \psi) \cup \{\varphi\}) \subseteq \Delta *_k (\psi \wedge \varphi)$.
   If φ is modally consistent with $\Delta *_k \psi$, then φ^τ is modally consistent with $\Delta *_a (\psi^\tau \wedge \mathscr{A}_L)$, and then by (K*8), $\mathrm{Cn}(\Delta *_a (\psi^\tau \wedge \mathscr{A}_L) \cup \{\varphi^\tau\}) \subseteq \mathrm{Cn}(\Delta^\tau *_a ((\psi \wedge \varphi)^\tau \wedge \mathscr{A}_L))$. But $\{\alpha \mid \Delta^\tau *_a (\psi^\tau \wedge \mathscr{A}_L) \vdash \alpha^\tau\} \cup \{\varphi\} \vdash_k \beta$ iff $\Delta^\tau *_a (\psi^\tau \wedge \mathscr{A}_L) \cup \{\varphi^\tau\} \vdash \beta^\tau$.

What we have just proven is that $*_k$ verifies all eight conditions of an AGM operation. Item 0c is actually a proof for Postulates (K*3) and (K*4).

The actual process of revision briefly discussed in the example above may be more complex than this. For instance, the object language might have a consistency notion other than that of classical logic or even none at all.

5. Revising Łukasiewicz's finitely many-valued logic $Ł_n$.

5.1. Łukasiewicz's finitely many-valued logic $Ł_n$.

5.1.1. Notation and terminology.

Syntax: Let $\mathscr{L}_{Ł_n}$ be the language of Łukasiewicz's finitely many-valued logic $Ł_n$ composed of a set of propositional letters $\mathscr{P}_{Ł_n} = \{p, q, \ldots\}$ and the connectives $\neg$ and $\rightarrow$. We use the special symbols $\top$ and $\bot$ to denote *truth* and *falsity*, respectively. The set of all well-formed formulae of $\mathscr{L}_{Ł_n}$, in symbols $WFF(\mathscr{L}_{Ł_n})$, is obtained by the standard construction mechanism of formulae. A theory presentation of $Ł_n$ is a (possibly infinite) set of well-formed formulae.

$Ł_n$ is characterised by the finite axiomatisation given below.

DEFINITION 5.1 (**Axiomatisation**). *Let α, β, and γ be formulae of $\mathscr{L}_{Ł_n}$. The following set of expressions together with the rule of modus ponens provide a proof theory for $Ł_n$. We refer to this set of axioms as $Ax_{Ł_n}$:*

https://doi.org/10.1017/S1755020308080246 Published online by Cambridge University Press

1. $\alpha \to \alpha$
2. $\alpha \to (\beta \to \alpha)$
3. $(\alpha \to \beta) \to ((\beta \to \gamma) \to (\alpha \to \gamma))$
4. $(\alpha \to \beta) \to ((\gamma \to \alpha) \to (\gamma \to \beta))$
5. $(\alpha \to (\beta \to \gamma)) \to (\beta \to (\alpha \to \gamma))$
6. $((\alpha \to \beta) \to \beta) \to ((\beta \to \alpha) \to \alpha)$
7. $((\alpha \to \beta) \to (\beta \to \alpha)) \to (\beta \to \alpha)$
8. $(\neg\beta \to \neg\alpha) \to (\alpha \to \beta)$
9. $((\alpha^n \to \beta) \to \alpha) \to \alpha$.

In the axiomatisation mentioned above, $\alpha^1 \to \beta \equiv \alpha \to \beta$, $\alpha^2 \to \beta \equiv \alpha \to (\alpha \to \beta)$, and so forth.

Semantics: We denote the set of n truth-values of Ł$_n$ by $\mathscr{T}_{Ł_n} = \left\{\frac{0}{n-1}, \frac{1}{n-1}, \ldots, \frac{n-1}{n-1}\right\}$, where $n \in \mathbb{N}$ and $n > 1$. We single out a set of designated values $\mathscr{D} = \left\{\frac{0}{n-1}, \ldots, \frac{j}{n-1}\right\}$, $0 \leq j < n-1$, and a set of undesignated values $\mathscr{U} = \left\{\frac{k}{n-1}, \ldots, \frac{n-1}{n-1}\right\}$, $k > j$. In this article, we assume that 0 is the only *designated* value and that 1 is the only *undesignated* value, that is, $j = 0$ and $k = 1$. A valuation of Ł$_n$ is a function $h : WFF(\mathscr{L}_{Ł_n}) \longrightarrow \mathscr{T}_{Ł_n}$, constructed from an initial assignment of truth-values to elements of $\mathscr{P}_{Ł_n}$ and extended to complex formulae by means of the truth-table given in Figure 1. Note that $\alpha \vee \beta$ can be defined as $(\alpha \to \beta) \to \beta)$ (this gives $s(\alpha \vee \beta) = \min(s(\alpha), s(\beta))$). We also use $\alpha \wedge \beta$ as a shorthand for $\neg((\neg A \to \neg B) \to \neg B)$ (this gives $s(\alpha \wedge \beta) = \max(s(\alpha), s(\beta))$).

Whenever $h(\varphi) < 1$, we say that the valuation h *satisfies* the formula φ, in symbols $h \Vdash_{Ł_n} \varphi$, or alternatively that the valuation h is a *model* of φ. We use the symbol $\models_{Ł_n} \varphi$ to denote the fact that $h(\varphi) = 0$, for all valuations h, that is, that φ is a tautology in Ł$_n$. Note that for all h, $h(\top) = 0$ and $h(\bot) = 1$.

We want to define what it means for a theory to be consistent. To do this, let us first look at an example. Remember that to us $1 = \bot$ and $0 = \top$.

Consider the formula $x \wedge \neg x$. Its truth-value is $\max(h(x), h(1-x)) > 0$. If $h(x) = 0$ or $h(x) = 1$, we get $h(x \wedge \neg x) = 1$. However, $x \wedge \neg x$ is *consistent* in the sense that for any value $0 < h(x) < 1$, we have the value of $h(x \wedge \neg x) < 1$.

Now consider the formula:

$$\neg(x \wedge \neg x).$$

This is not a tautology because its truth-value is

$$1 - h(x \wedge \neg x),$$

and we already showed that for $0 < h(x) < 1$, $h(x \wedge \neg x) < 1$, and hence, for $0 < h(x) < 1$, $h(\neg(x \wedge \neg x)) > 0$. To be a tautology, a formula must get 0 for all possible assignments.

Note that $h(\neg\varphi) = 1 - h(\varphi) = \max(0, 1 - h(\varphi)) = h(\varphi \to \bot)$. We take the notion of consistency of $\Delta = \{A_1, \ldots, A_n\}$ to say that Δ does not entail $\bot$ (i.e., that $\bot$ is not a consequence of Δ). But what does 'entail' mean? There are two possibilities in the literature for

$$A_1, \ldots, A_n \vDash B.$$

$s(\varphi)$	$s(\psi)$	$s(\neg\varphi)$	$s(\varphi \to \psi)$	$s(\varphi \wedge \psi)$	$s(\alpha \vee \beta)$
x	y	$1-x$	$\max(0, y-x)$	$\max(x, y)$	$\min(x, y)$

Fig. 1. Truth-tables for the connectives in Ł$_n$.

https://doi.org/10.1017/S1755020308080246 Published online by Cambridge University Press

We denote them by $\models_{\max}$ and $\models_{\text{sum}}$. We start with the latter. We want that $A_1, \ldots, A_n \models_{\text{sum}} B$ if and only if $\models A_1 \to (A_1 \to \ldots \to (A_n \to B)\ldots)$. In other words, we want the implication $A_1 \to (A_1 \to \ldots \to (A_n \to B)\ldots)$ to be a tautology. For this to happen, we need that for any h, $h(A_1 \to \ldots \to (A_n \to B)\ldots)) = 0$ if and only if $h(A_1) + \ldots + h(A_n) = \sum_{i=1}^{n} h(A_i) \geq h(B)$. In general, we have the following.

DEFINITION 5.2 ($\models_{\text{sum}}$). *Let Δ be a set of formulae and φ a formula of $\mathscr{L}_{Ł_n}$.*

$$\Delta \models_{\text{sum}} B \text{ if and only if } \sum_{A_i \in \Delta} h(A_i) \geq h(B).$$

The other definition of $A_1, \ldots, A_n \models B$ is the 'max' definition. We want $A_1, \ldots, A_n \models_{\max} B$ if and only if $\models A_1 \wedge \ldots \wedge A_n \to B$. In semantical terms, this means that $A_1, \ldots, A_n \models_{\max} B$ if and only if for all h, we have $\max_{i=1}^{n}\{h(A_i)\} \geq h(B)$. More generally, we have the following.

DEFINITION 5.3 ($\models_{\max}$). *Let Δ be a set of formulae and φ a formula of $\mathscr{L}_{Ł_n}$.*

$$\Delta \models_{\max} B \text{ if and only if } \max_{A_i \in \Delta}\{h(A_i)\} \geq h(B).$$

REMARK 5.4. *Note that we are not saying that*

$$A_1, \ldots, A_n \models'_{\max} B$$

if every model of $\{A_i\}$ is also a model of B. This would mean that for every h, if $\max h(A_i) < 1$, then $h(B) < 1$.

According to this definition, we get:

$$\top \models'_{\max} (A \to \neg A) \to \neg A.$$

because for every h, $t = h((A \to \neg A) \to \neg A) < 1$ since $t = \max(0, 1 - h(A) - \max(0, 1-h(A)-h(A))) = \max(0, 1-h(A)-\max(0, 1-2h(A)))$. If $h(A) = 0$, $t = 0 < 1$. For $0 < h(A) < \frac{1}{2}$, $t = h(A) < 1$. For $\frac{1}{2} < h(A) \leq 1$, we have $t = 1 - h(A) < 1$.

We see that when we translate into classical logic, we get this $\models'_{\max}$ consequence relation.

We prefer the $\models_{\max}$ definition. Note that the $\models_{\text{sum}}$ definition satisfies the deduction theorem

$$A_1, \ldots, A_n \models_{\text{sum}} B \to C \text{ if and only if } A_1, \ldots, A_n, B \models_{\text{sum}} C.$$

But its weakness from our point of view is that it regards the theory as a multi-set of a *sequence*, not as a *set*.

We have, for example, in three-valued logic (with truth-values $\{0, \frac{1}{2}, 1\}$) that $x \wedge \neg x \not\models_{\text{sum}} \bot$ but $x \wedge \neg x, x \wedge \neg x \models_{\text{sum}} \bot$. This is because

$$x \wedge \neg x \to \bot$$

is not a tautology but

$$(x \wedge \neg x) \to ((x \wedge \neg x) \to \bot)$$

is a tautology.

So, if we look at $\Delta = \{x \wedge \neg x\}$ as a set, then we can certainly take as many copies of its members in a proof as we want. Thus, the consequence relation $\models_{\text{sum}}$ regards Łukasiewicz's logic as a sub-structural logic where theories are multi-sets. Let us hence concentrate on the $\models_{\max}$ consequence notion.

https://doi.org/10.1017/S1755020308080246 Published online by Cambridge University Press

The reader is warned that $\models_{\max}$ can have unexpected results. Consider

$$\{A, A \to B\} \overset{?}{\models} B.$$

According to $\models_{\text{sum}}$, we have to check whether $(A \to B) \to (A \to B)$ is a tautology or alternatively check whether $A \to ((A \to B) \to B)$ is a tautology.

So indeed, $\{A, A \to B\} \models_{\text{sum}} B$. How about

$$\{A, A \to B\} \overset{?}{\models}_{\max} B.$$

We have to check whether $(A \wedge (A \to B)) \to B$ is a tautology. Let $h(A) = \frac{1}{5}$, $h(B) = \frac{3}{5}$. Hence,

$$h(A \to B) = \tfrac{2}{5}$$
$$h(A \wedge (A \to B)) = \tfrac{2}{5},$$

and therefore

$$h((A \wedge (A \to B)) \to B) = \frac{1}{5}.$$

Thus, $\{A, A \to B\} \not\models_{\max} B$. We do have, however, for any single formula A that

$$A \models_{\max} B \text{ if and only if } A \models_{\text{sum}} B.$$

Note that this does not hold for $\models'_{\max}$ because as we have showed

$$\top \models'_{\max} (A \to \neg A) \to \neg A$$

but not

$$\models'_{\max} \top \to ((A \to \neg A) \to \neg A).$$

$\models'_{\max}$ takes as tautologies anything that *never* gets $\bot$.

So we take $\models_{\max}$ as our consequence notion.

What does it means now for a theory to be $\models_{\max}$ -consistent? It means that it does not have $\bot$ as a consequence.

DEFINITION 5.5 (**Weak consistency**). *A theory Δ is weakly consistent if for some assignment h,*

$$h(\Delta) = \max\{h(A) | A \in \Delta\} < 1.^4$$

Analogously, we can define a stronger notion of consistency.

DEFINITION 5.6 (**Strong consistency**). *A theory Δ is strongly consistent if for some assignment h, $h(\Delta) \in \mathscr{D}$.*

The difference becomes clearer when we want to close a theory Δ under the consequence relation. For comparison, for the case of $\models_{\text{sum}}$, we define the notion of weak consistency for a multi-set Δ by letting,

$$h(\Delta) = \sum_{A \in \Delta} h(A),$$

and say that Δ is consistent if and only if there exists an assignment h such that $h(\Delta) < 1$. So, for example, in Ł$_2$, $\{x \wedge \neg x\}$ is consistent, while $\{x \wedge \neg x, x \wedge \neg x\}$ is not.

We are now in a position to define the notion of consequence closure of a theory Δ.

---

[4] In general, Δ is consistent if for some h, $h(\Delta) \notin \mathscr{U}$.

https://doi.org/10.1017/S1755020308080246 Published online by Cambridge University Press

DEFINITION 5.7 (**Consequence closure for** $\models_{\max}$). *Let* Δ *be a theory.*

$$\mathrm{Cn}_{\max}(\Delta) = \{B | \textit{ for some finite } \Delta' \subseteq \Delta, \textit{ we have that } \models \wedge_{\varphi_i \in \Delta'} \varphi_i \rightarrow B\}.$$

LEMMA 5.8. $\mathrm{Cn}_{\max}(\mathrm{Cn}_{\max}(\Delta)) = \mathrm{Cn}_{\max}(\Delta)$.

Proof. Assume that

1. $\theta_i \subseteq \Delta, i = 1, \ldots, m$
2. $\models \wedge\theta_i \rightarrow B_i, i = 1, \ldots, m$
3. $\models \wedge B_i \rightarrow C$.

Parts (1) and (2) say that $B_i \in \mathrm{Cn}_{\max}(\Delta)$. Part (3) says that $C \in \mathrm{Cn}_{\max}(\mathrm{Cn}_{\max}(\Delta))$. We want to find a $\theta \subseteq \Delta$ such that $\models \wedge\theta \rightarrow C$. This will show that $C \in \mathrm{Cn}_{\max}(\Delta)$. Let $\theta = \bigcup_i \theta_i$. If $\not\models \wedge\theta \rightarrow C$, then for h, $h(\theta) < h(C)$. However, from Part (1), $\max\{h(\theta_i)\} \geq h(B_i)$, and from Part (3), $\max\{h(B_i)\} \geq h(C)$. Let B_j be the max given. Then we have,

$$h(\theta) \geq h(\theta_j) \geq h(B_j) \geq h(C),$$

a contradiction. The other inclusion can be proved in a straightforward way. □

LEMMA 5.9. $\mathrm{Cn}_{\max}(\mathrm{Cn}_{\max}(\Delta) \cup \{\varphi\}) = \mathrm{Cn}_{\max}(\Delta \cup \{\varphi\})$.

Proof. The proof is similar to the proof of Lemma 5.8, except that in it, we need to consider the set $\Delta \cup \{\varphi\}$ instead of Δ. □

REMARK 5.10. *As a matter of interest, how do we define consequence closure for the case of* $\models_{\mathrm{sum}}$*?*

We can try to give a similar definition.

$$\mathrm{Cn}_{\mathrm{sum}}(\Delta) = \{B | \text{ for some } A_1, \ldots, A_n \text{ in } \Delta, \models A_1 \rightarrow \ldots \rightarrow (A_n \rightarrow B)\ldots)\}.$$

But this is not so simple, Δ being a multi-set. We let

$$\Delta = \{A\}.$$

So $\mathrm{Cn}_{\mathrm{sum}}(\Delta) = \{A, A\}$,

$$\mathrm{Cn}_{\mathrm{sum}}(\mathrm{Cn}_{\mathrm{sum}}(\Delta)) = \{A, A, A, A\},$$

and so forth. We saw that $\{x \wedge \neg x\}$ is consistent, while $\{x \wedge \neg x, x \wedge \neg x\}$ is not. See next lemma for $\models_{\max}$.

LEMMA 5.11 (**for** $\models_{\max}$).

1. $\Delta \subseteq \mathrm{Cn}_{\max}(\Delta)$.
2. *If* Δ *is weakly consistent, so is* $\mathrm{Cn}_{\max}(\Delta)$.

Proof.

1. $A \in \Delta$, then $A \in \mathrm{Cn}_{\max}(\Delta)$ because $\Delta \models_{\max} A \rightarrow A$.
2. If $h(\Delta) < 1$ and $\theta \subseteq \Delta$ and $\models \wedge\theta \rightarrow B$, then $h(\theta) \geq h(B)$, and since $h(\Delta) \geq h(\theta)$, we get $h(\Delta) \geq h(B)$. But this means that $h(\mathrm{Cn}_{\max}(\Delta)) = h(\Delta) < 1$. □

LEMMA 5.12. *Let* Δ *be a theory and let* $B \notin \mathrm{Cn}_{\max}(\Delta)$ *(i.e.,* $\Delta \not\models_{\max} B$*). Then for some h,*

$$h(\Delta) < h(B).$$

https://doi.org/10.1017/S1755020308080246 Published online by Cambridge University Press

Proof. If $\Delta \not\models_{\max} B$, then for any $\theta \subseteq \Delta, \theta$ finite, we have $\not\models_{\max} \wedge\theta \to B$. So, for some assignment h_θ such that $h_\theta(\theta \to B) > 0$. Therefore, if $\theta = \{A_i\}$,

$$\max\{h_\theta(A_i)\} < h_\theta(B).$$

We want a h_Δ such that

$$\max\{h_\Delta(A) | A \in \Delta\} < h_\Delta(B).$$

For each $\theta \subseteq \Delta, \theta$ finite, let $Q_\theta = \{h | h(\theta \to B) > 0\}$. Clearly by our assumptions, Q_θ is not empty. Also, $Q_\theta \cap Q_{\theta'} = Q_{\theta\cup\theta'} \neq \varnothing$. For we have for any h,

$$\max\{h(A) | A \in \theta\} < h(B), \text{ and } \max\{h(A) | A \in \theta'\} < h(B)$$
$$\text{if and only if}$$
$$\max\{h(A) | A \in \theta \cup \theta'\} < h(B).$$

Thus, the family Q_θ is a family of non-empty sets closed under intersection. This family can be extended to an ultrafilter $\mathscr{F}$. Thus, $Q_\theta \in \mathscr{F}$ for all $\theta \subseteq \Delta, \theta$ finite.

Let q be any atom. Consider the sets:

$$S_q^x = \{h | h(q) = x\} \text{ for } x \in \left\{0, \frac{1}{n-1}, \ldots, \frac{n-1}{n-1}\right\}.$$

These sets are disjoint, so only one of them is a large set. Let h_Δ be such that

$$h_\Delta(q) = x \text{ if and only if } S_q^x \in \mathscr{F}.$$

We thus get an assignment to the atoms of the language. We claim

$$\max\{h_\Delta(A) | A \in \Delta\} < h_\Delta(B).$$

Let $S_\varphi^x = \{h | h(\varphi) = x\}$. □

LEMMA 5.13. *For all* φ,

$$S_\varphi^{H_\Delta(\varphi)} \in \mathscr{F}.$$

Proof. By induction on φ. For atomic q, this is the definition. Assume that this holds for φ_1 and φ_2, and let $h_\Delta(\varphi_1) = x_1$ and $h_\Delta(\varphi_2) = x_2$. So,

$$y = h_\Delta(\varphi_1 \to \varphi_2) = \max(0, h_\Delta(\varphi_2) - h_\Delta(\varphi_1)) = \max(0, x_2 - x_1).$$

So the set of all h such that $h(\varphi_i = x_i)$ is a large set in $\mathscr{F}$, and hence, the intersection is also in $\mathscr{F}$. But this is the set of h that gives $h(\varphi_1) = x_1$ and $h(\varphi_2) = x_2$, and they give $h(\varphi_1 \to \varphi_2) = y$. So $S_\varphi^y \in \mathscr{F}$.

Assume now that for some $A \in \Delta$,

$$h_\Delta(A \to B) = 0.$$

Then, $S_{A\to B}^0 \in \mathscr{F}$. But also $S_A \in \mathscr{F}$, being the set of all h such that $h(A \to B) > 0$.

This is impossible because any two sets in $\mathscr{F}$ must have a non-empty intersection. □

5.2. Translating Łukasiewicz's many-valued logic $Ł_n$ into classical logic. We assume that the language of classical logic is formulated with the connectives $\vee$, $\wedge$, $\neg$, and $\to$, and that its atoms are $\mathscr{P}_n^* = \left\{p_0^{*,n}, \ldots, p_{\frac{1}{n-1}}^{*,n}, \ldots, p_{\frac{n-1}{n-1}}^{*,n}, q_0^{*,n}, \ldots, q_{\frac{1}{n-1}}^{*,n}, \ldots, q_{\frac{n-1}{n-1}}^{*,n}, \ldots\right\}$.

A valuation h^* of classical logic gives values to the atoms in $\mathscr{P}_n^*$ in $\{\top, \bot\}$ and is extended to complex formulae in the usual way. We use the symbol $\models_{*,n} B^{*,n}$ to denote the fact that the formulae $B^{*,n}$ is a tautology in classical logic.

https://doi.org/10.1017/S1755020308080246 Published online by Cambridge University Press

To lighten the notation, we drop the superscript n from now on since we assume an arbitrary but fixed number of truth-values in $Ł_n$.

5.2.1. Agreement for translation from $Ł_n$ to classical logic. Our translation is done using the following correspondence:

$$h(p) = \frac{i}{n-1} \text{ if and only if } h^*\left(p^*_{\frac{i}{n-1}}\right) = \top \text{ and } h^*\left(p^*_{\frac{j}{n-1}}\right) = \bot \text{ for } j \in \{0,\ldots,n-1\} - \{i\}.$$

In other words, p^*_x is true in a classical logic valuation h^* exactly when the truth-value of p in $Ł_n$ is x in some valuation h and the truth-values of all other p^*_y for $y \neq x$ in h^* is false.

For that, we need regulating formulae $\Psi^*_{p_i}$ of classical logic for each $p_i \in \mathscr{P}_{Ł_n}$:

$$\Psi^*_{p_i} = \left[\bigvee_{x \in \mathscr{T}_{Ł_n}} (p_i)^*_x\right] \wedge \bigwedge_{x,y \in \mathscr{T}_{Ł_n}, x \neq y} \neg((p_i)^*_x \wedge (p_i)^*_y).$$

The formula $\Psi^*_{p_i}$ is satisfied by a valuation h^* only when exactly one of $p^*_{\frac{i}{n-1}}$ (for $i = 0, \ldots, n-1$) gets value $\top$ and the others get $\bot$ and hence regulates theories of classical logic containing it to behave as a particular assignment h of $Ł_n$.

DEFINITION 5.14. *Let φ be a formula in $\mathscr{L}_{Ł_n}$. We use the term $atoms(\varphi)$ to denote the set of atoms $\{p, q, r, \ldots\}$ appearing in φ. Analogously, we use the term $atoms^*(\varphi)$ to denote the corresponding atoms $\left\{p^*_0, \ldots, p^*_{\frac{n-1}{n-1}}, q^*_0, \ldots, q^*_{\frac{n-1}{n-1}}, \ldots\right\}$ in classical logic.*

DEFINITION 5.15. *Let φ be a formula in $\mathscr{L}_{Ł_n}$. $\tau_n(\varphi) = \wedge_{p_i \in atoms(\varphi)} \Psi^*_{p_i}$.*

We now offer a translation $(*)$ of $Ł_n$ into classical logic. We start with the statement of $Ł_n$ saying that 'the value of the formula A in a valuation h is x' (which we denote by $[A]_x$).

DEFINITION 5.16 (**The translation function** $*$).

1. *For $p \in \mathscr{P}_{Ł_n}$, let $[p]^*_x = p^*_x$*
2. $[\neg A]^*_x = [A]^*_{1-x}$
3. $[A \to B]^*_x = \bigvee_{y,z \in \mathscr{T}_{Ł_n} s.t. x = \max(0, z-y)} [A]^*_y \wedge [B]^*_z$
4. $\Delta^*_x = \{\varphi^*_x \mid \varphi \in \Delta\}$.

The function $*$ captures in first-order logic the truth-value of a formulae of $\mathscr{L}_{Ł_n}$ in a particular assignment h. Later on, we use this definition in the translation according to a particular entailment semantics (e.g., $\vDash_{\max}$).

DEFINITION 5.17. *Let φ be a formula in $\mathscr{L}_{Ł_n}$. $\pi_n(\varphi) = \bigvee_{x \in \mathscr{T}_{Ł_n}} [\varphi]^*_x \wedge \bigwedge_{x,y \in \mathscr{T}_{Ł_n}, x \neq y} \neg$ $([\varphi]^*_x \wedge [\varphi]^*_y)$.*

LEMMA 5.18. *Let φ be a formula in $\mathscr{L}_{Ł_n}$; $\tau_n(\varphi)$ as given in Definition 5.15 and $\pi_n(\varphi)$ as given in Definition 5.17. It follows that*

$$\vDash_* \tau_n(\varphi) \to \pi_n(\varphi).$$

Proof. The proof is done by induction on φ.

1. For φ atomic, that is, $\varphi = p$ for some $p \in \mathscr{P}_{Ł_n}$, this holds because $\tau_n(p) = \pi_n(p)$.

https://doi.org/10.1017/S1755020308080246 Published online by Cambridge University Press

2. Assume $\pi_n(\varphi)$ holds. We show that $\neg\pi_n(\varphi)$ also holds. We know that

$$\tau_n \models_* \left(\bigvee_{x\in\mathscr{T}_{Ł_n}} [\varphi]^*_x\right) \wedge \bigwedge_{x,y\in\mathscr{T}_{Ł_n}, x\neq y} \neg([\varphi]^*_x \wedge [\varphi]^*_y).$$

We need to show that the following holds:

$$\left(\bigvee_{x\in\mathscr{T}_{Ł_n}} [\neg\varphi]^*_x\right) \wedge \bigwedge_{x,y\in\mathscr{T}_{Ł_n}, x\neq y} \neg([\neg\varphi]^*_x \wedge [\neg\varphi]^*_y).$$

But notice that $[\neg\varphi]^*_u = [\varphi]^*_{1-u}$ and since u ranges over $\mathscr{T}_{Ł_n}$ so does $1-u$. Hence, the first conjunct holds and so does the second one for the same reason.

3. Consider now $\varphi = A \to B$. We have:

$$[\varphi]^*_x = \bigvee_{y,z\in\mathscr{T}_{Ł_n} \text{ s.t. } x=\max(0,z-y)} [A]^*_y \wedge [B]^*_z.$$

Let h^* be any classical logic model of τ_n. By the induction hypothesis under τ_n in classical logic, there exists a unique y and a unique z such that $[A]^*_y$ and $[B]^*_z$ hold. This is because $\pi_n(A)$ and $\pi_n(B)$ hold. If we pick $x = \max(0, z-y)$, we then have that $[\varphi]^*_x$ holds. Therefore, $\bigvee_{x\in\mathscr{T}_{Ł_n}} [\varphi]^*_x$ holds in any model of τ_n. We now need to show that $\bigwedge_{x_1,x_2\in\mathscr{T}_{Ł_n}, x_1\neq x_2} \neg([A\to B]^*_{x_1} \wedge [A\to B]^*_{x_2})$ holds in the model. So assume that for some x_1 and x_2, we have that $[A\to B]^*_{x_1}$ and $[A\to B]^*_{x_2}$ hold in the model. But since by the induction hypothesis there are unique values y and z such that $[A]^*_y$ and $[B]^*_z$ hold, we get that there is a unique $x = \max(0, z-y)$ such that $[A\to B]^*_x$ holds. Therefore, $x = x_1 = x_2$.

□

What the lemma above says is that any model h^* of a formula $\tau_n(\varphi)$ is also a model of $\pi_n(\varphi)$, and hence, we can work with the constraint $\tau_n(\varphi)$ imposed on the atoms of φ only.

We now extend τ_n to a set of formulae Δ in the following way: $\tau_n(\Delta) = \{\tau_n(\varphi) \mid \varphi \in \Delta\}$.

LEMMA 5.19. *Let h be a valuation of $Ł_n$ and h^* be defined in classical logic by correspondence (#) below.*

(#) $h^*(q^*_x) = \top$ *if and only if* $h(q) = x$

It follows that $h^* \models_* \tau_n(\Delta)$ *and* $h^* \models_* [\varphi]^*_{h(\varphi)}$ *for all* $\varphi \in \Delta$.

Proof. The proof is done by induction on φ using Lemma 5.18. □

Now since the models of a formula φ were defined as the valuations that assign to it values that are not undesignated, we complete the translation to classical logic in the following way.

5.2.2. Translating into classical logic (case of $\models_{\max}$). Given a theory Δ, consider $\mathrm{Cn}_{\max}(\Delta)$, the consequence closure of Δ. It being consistent means for some h, $h(A) < 1$ for all $A \in \mathrm{Cn}_{\max}(\Delta)$.

NOTATION 5.20. *Let φ be a formula of $\mathscr{L}_{Ł_n}$. The term $[\varphi]^*_{<1}$ will abbreviate the formula $\bigvee_{x<1}[\varphi]^*_x$.*

DEFINITION 5.21. *Let φ be a formula in $\mathscr{L}_{Ł_n}$. The translation of φ into classical logic, in symbols φ^τ, is the formula $\tau_n(\varphi) \wedge [\varphi]^*_{<1}$. Similarly, for a set of formulae Δ, we define $\Delta^\tau = \{\varphi^\tau \mid \varphi \in \Delta\}$.*

https://doi.org/10.1017/S1755020308080246 Published online by Cambridge University Press

Then, $\mathrm{Cn}_{\max}(\Delta)$ is consistent if and only if $\mathrm{Cn}_{\max}(\Delta)^{\tau}$ is consistent in classical logic.

LEMMA 5.22 (**Correspondence of consistency**). *Let Δ be a set of formulae of $\mathscr{L}_{Ł_n}$ and φ and ψ formulae of $\mathscr{L}_{Ł_n}$. It follows that:*

1. *φ is consistent in $Ł_n$ if and only if φ^{τ} is consistent in classical logic.*
2. *Δ is consistent in $Ł_n$ if and only if Δ^{τ} is consistent in classical logic.*

Proof. The proof of Part (1) follows from Lemma 5.19 and the proof of Part (2) given Part (1) is straightforward. □

Notice that we do not need to define a notion of acceptability for the revision of a theory of $Ł_n$, since inconsistency in $Ł_n$ translates directly into inconsistency into classical logic.

EXAMPLE 5.23. *Let $\varphi = \neg(p \to p)$, which is inconsistent in $Ł_n$ since $h(\neg(p \to p)) = 1$ for all h. We only need to show that the translation will include $[p]^*_z$ and $[p]^*_y$ for some z, y, $z \neq y$ since this contradicts $\tau_n(\varphi)$, which says that the truth-values of propositional variables in φ are uniquely assigned.*

Note that each disjunct in $\bigvee_{x<1}[\neg(p \to p)]^*_x$ will include $[p \to p]^*_{1-x} = \bigvee_{z>y}([p]^*_y \wedge [p]^*_z)$. This is because since $x < 1$, then $1 - x > 0$ and hence $1 - x = \max(0, z - y) > 0$ and $z > y$. Provided there are at least two values in $\mathscr{T}_{Ł_n}$, there is at least one such pair of z and y, and this contradicts $\tau_n(\varphi)$.

5.3. Revision in Łukasiewicz's many-valued logic ($Ł_n$). Revision in $Ł_n$ is defined below as a two-stage process.

DEFINITION 5.24 (**Revision by translation of** $Ł_n$). *Let Δ be a set of formulae in $\mathscr{L}_{Ł_n}$ and φ a formula in $\mathscr{L}_{Ł_n}$.*

Let the set $\Delta^{\#}_{\varphi}$ be defined as:

$$\Delta^{\#}_{\varphi} = \{\psi \in \mathrm{Cn}_{\max}(\Delta) \mid [\psi]^*_{<1} \in \mathrm{Cn}(\mathrm{Cn}_{\max}(\Delta)^*_{<1}) *_a (\tau_n(\Delta) \wedge \varphi^{\tau})\}.$$

The revision of Δ by φ, in symbols $\Delta *_{Ł_n} \varphi$, is defined as follows.

$$\Delta *_{Ł_n} \varphi = \mathrm{Cn}_{\max}(\Delta^{\#}_{\varphi} \cup \{\varphi\}).$$

Note that with respect to the general methodology presented in Section 3, the axiomatisation $\mathscr{A}_L$ is for this logic embedded into the constraints $\tau_n(\Delta)$ and $\tau_n(\varphi)$, whereas the consistency notion is captured by the constraints $[\alpha]^*_{<1}$, for all formulae α involved in the process.

Furthermore, the following points need to be taken into account.

1. We first take the set of consequences of Δ under $\models_{\max}$.
2. Then we need to translate every formulae in this set to classical logic (stating that they can have any non-undesignated value).

We now need to work on the revising part. We need to include the following:

1. The constraints on the behaviour of the truth-values for the atoms appearing in Δ (i.e., $\tau_n(\Delta)$), which is included in this part to make sure it is preserved during the revision.
2. The translation of the revising formula φ itself: φ^{τ}.

https://doi.org/10.1017/S1755020308080246 Published online by Cambridge University Press

Unfortunately, we are not done yet, as the classical logic revision will return many more consequences than the ones that were originally in $\mathrm{Cn}_{\max}(\Delta)$. Some of the new ones are legitimate since they are the result of the addition of φ to Δ; however, some will in fact be $\max'$-consequences of Δ, for example, $A \vee \neg A$ since the translation of $A \vee \neg A$ gets value less than 1. To single them out, we maintain only those formulae ψ that were originally in $\mathrm{Cn}_{\max}(\Delta)$. This is defined in $\Delta^{\#}_{\varphi}$. Note that $\Delta^{\#}_{\varphi}$ excludes φ and its consequences in all interesting cases; therefore, we need to add it close the result under $\mathrm{Cn}_{\max}$.

We could have defined the revision by working with the subsets of $\mathrm{Cn}_{\max}(\Delta)$ that are maximally consistent with φ, but we would have the problem of choosing the right subset of this set. If we want to mimic the behaviour of the $*_a$ operation, we let it do its job and weed out the consequences unwantedly added by the classical logical consequence closure.

LEMMA 5.25. *Let Δ be a theory of Ł$_n$, and φ and ψ formulae of $\mathscr{L}_{Ł_n}$. The following properties hold for $*_{Ł_n}$.*

1. *$\Delta *_{Ł_n} \varphi$ is closed under logical consequence.*
   This comes straight from Definition 5.24.
2. *$\Delta *_{Ł_n} \varphi \models_{Ł_n} \varphi$.*
   Note that according to Definition 5.24, $\varphi \in \Delta *_{Ł_n} \varphi$. Hence, $\max_{A_i \in \Delta *_{Ł_n} \varphi}\{h(A_i)\} \geq h(\varphi)$, and then by Definition 5.3, $\Delta *_{Ł_n} \varphi \models_{Ł_n} \varphi$.
3. *If φ is consistent with Δ, then $\Delta *_{Ł_n} \varphi = \mathrm{Cn}_{\max}(\Delta \cup \{\varphi\})$.*
   *We already know from Lemma 5.22 that if $\Delta \cup \{\varphi\}$ is consistent, then so is the set $\Delta^{\tau} \cup \{\varphi^{\tau}\}$ in classical logic. Consequently, so is the set $\mathrm{Cn}(\mathrm{Cn}_{\max}(\Delta)^*_{<1}) \cup \{(\tau_n(\Delta) \wedge \varphi^{\tau})\}$ and hence by (K*3, 4), $\mathrm{Cn}(\mathrm{Cn}_{\max}(\Delta)^*_{<1}) * \{(\tau_n(\Delta) \wedge \varphi^{\tau})\} = \mathrm{Cn}(\mathrm{Cn}_{\max}(\Delta)^*_{<1}) \cup \{(\tau_n(\Delta) \wedge \varphi^{\tau})\}$. Note that by the way we have defined the translation*

$$\{\psi \in \mathrm{Cn}_{\max}(\Delta) \mid [\psi]^*_{<1} \in \mathrm{Cn}(\mathrm{Cn}_{\max}(\Delta)^*_{<1}) \cup \{(\tau_n(\Delta) \wedge \varphi^{\tau})\} = \mathrm{Cn}_{\max}(\Delta).$$

   By Lemma 5.9, $\mathrm{Cn}_{\max}(\mathrm{Cn}_{\max}(\Delta) \cup \{\varphi\}) = \mathrm{Cn}_{\max}(\Delta \cup \{\varphi\})$.
4. *$\Delta *_{Ł_n} \varphi$ is inconsistent, only if φ is contradictory.*
   *By (K*5), Lemma 5.22, and Definition 5.24, $\Delta^{\#}_{\varphi}$ is inconsistent only if φ^{τ} is inconsistent, only if φ is contradictory. Therefore, $\Delta *_{Ł_n} \varphi = \mathrm{Cn}_{\max}(\Delta^{\#}_{\varphi} \cup \{\varphi\})$ is inconsistent only if φ is contradictory.*
5. *If $\varphi \equiv_{Ł_n} \psi$, then $\Delta *_{Ł_n} \varphi \equiv \Delta *_{Ł_n} \psi$.*
   *If $\varphi \equiv_{\max} \psi$, then for all h, $h(\varphi) = h(\psi)$. Therefore, $\varphi^{\tau} \equiv \psi^{\tau}$. By (K*6), $\mathrm{Cn}(\mathrm{Cn}_{\max}(\Delta)^*_{<1}) * \{(\tau_n(\Delta) \wedge \varphi^{\tau})\} \equiv \mathrm{Cn}(\mathrm{Cn}_{\max}(\Delta)^*_{<1}) * \{(\tau_n(\Delta) \wedge \psi^{\tau})\}$ and hence $\Delta^{\#}_{\varphi} = \Delta^{\#}_{\psi}$. Since $\varphi \equiv_{Ł_n} \psi$, we have that $\Delta *_{Ł_n} \varphi \equiv \Delta *_{Ł_n} \psi$.*
6. *$\Delta *_{Ł_n} (\varphi \wedge \psi) \subseteq \mathrm{Cn}_{Ł_n}((\Delta *_{Ł_n} \varphi) \cup \{\psi\})$.*
   *Suppose that $\alpha \in \Delta *_{Ł_n} (\varphi \wedge \psi)$, for some α. By Definition 5.24, $\alpha \in \mathrm{Cn}_{\max}(\Delta^{\#}_{\varphi \wedge \psi} \cup \{\varphi \wedge \psi\}$. By (K*7), $\mathrm{Cn}(\mathrm{Cn}_{\max}(\Delta)^*_{<1}) *_a (\tau_n(\Delta) \wedge (\varphi \wedge \psi)^{\tau}) \subseteq \mathrm{Cn}(\mathrm{Cn}_{\max}(\Delta)^*_{<1}) *_a (\tau_n(\Delta) \wedge (\varphi)^{\tau}) \cup \{\psi\})$*
   *$\Delta^*_0 *_a (\tau_n(\Delta) \wedge \varphi^{\tau} \wedge \psi^{\tau}) \vdash \alpha^{\tau}$. By (K*7), it follows that $\mathrm{Cn}(\Delta^*_0 *_a (\tau_n(\Delta) \wedge \varphi^{\tau} \cup \{\psi^{\tau}\} \vdash \alpha^{\tau}$. Notice that for every $\psi^{\tau} \in \Delta^*_0 *_a (\tau_n(\Delta) \wedge \varphi^{\tau}$, there is a corresponding ψ in $\Delta *_{Ł_n} \varphi$ (by the definition of $*_{Ł_n}$) and similarly for ψ^{τ}. By correspondence, $\mathrm{Cn}_{Ł_n}((\Delta *_{Ł_n} \varphi) \cup \{\psi\}) \vdash \alpha$.*
7. *If ψ is consistent with $\Delta *_{Ł_n} \varphi$, then $\mathrm{Cn}((\Delta *_{Ł_n} \varphi) \cup \{\psi\}) \subseteq \Delta *_{Ł_n} (\varphi \wedge \psi)$.*
   *If ψ is consistent with $\Delta *_{Ł_n} \varphi$, then ψ^{τ} is consistent with $\Delta^*_0 *_a (\tau_n(\Delta) \wedge \varphi^{\tau})$, and then by (K*8), $\mathrm{Cn}(\Delta^*_0 *_a (\tau_n(\Delta) \wedge \varphi^{\tau}) \cup \{\psi^{\tau}\}) \subseteq \mathrm{Cn}(\Delta^*_0 *_a (\tau_n(\Delta) \wedge (\varphi \wedge \psi)^{\tau})$.*

https://doi.org/10.1017/S1755020308080246 Published online by Cambridge University Press

But $\{\alpha \mid \Delta_0^* *_a(\tau_n(\Delta) \wedge \varphi^\tau) \vdash \alpha^\tau\} \cup \{\psi\} \models_{Ł_n} \beta$ *if and only if* $\Delta_0^* *_a(\tau_n(\Delta) \wedge \varphi^\tau) \cup \{\psi^\tau\} \vdash \beta^\tau$.

What we have just proven is that $*_{Ł_n}$ *verifies all eight conditions of an AGM operation. Item 3 is actually a proof for Postulates* (K*3) *and* (K*4).

6. Revising algebraic logics.

6.1. Translating algebraic logic into classical logic. We now present a translation method for a class of algebraic logics. This class includes Łukasiewicz's infinitely many-valued logic and classical logic itself.

Let $\mathfrak{L}$ be an algebraic logic defined as follows.

1. $\mathfrak{L}$ has connectives $c_1, \ldots, c_k$, where each c_i is an m_i-ary connective.
2. $\mathfrak{L}$ has as truth-values elements of any algebra of the form $\mathfrak{A} = (\mathscr{M}, \mathfrak{D}, f_1, \ldots, f_k)$, where $\mathscr{M}$ is a domain of algebraic elements; $\mathfrak{D} \subseteq \mathscr{M}$ is the set of designated elements; and each f_i is a m_i-ary function symbol corresponding to the connectives $c_1, \ldots, c_k$.

We assume that the class of models that can serve $\mathfrak{L}$ is defined by some first-order theory Δ. We can assume Δ is finite.

Let $\mathfrak{A}$ be an algebra. A valuation h into $\mathfrak{A}$ assigns an algebraic element $h(q) \in \mathscr{M}$ to every atom q of $\mathfrak{L}$. h is extended to any formula φ of the language by recursively defining

$$h(c_i(A_1, \ldots, A_{m_i})) = f_i(h(A_1), \ldots, h(A_{m_i})),$$

where c_i is the main connective in φ and the A_{m_i}s are the arguments of c_i.

We say that $\mathfrak{L}_\mathfrak{A} \models \varphi$ if and only if for all valuations h into $\mathfrak{A}$, we have that $h(\varphi) \in \mathfrak{D}$. We say that $\mathfrak{L}_\Delta \models \varphi$ if and only if for any model $\mathfrak{A}$ of Δ, we have that $\mathfrak{L}_\mathfrak{A} \models \varphi$.

EXAMPLE 6.1.

1. *Classical propositional logic is an algebraic logic for an axiomatic theory of Boolean algebras* Δ.
2. *Łukasiewicz' infinitely many-valued propositional logic is an algebraic logic for the theory of MV-algebras* Δ.
3. *Intuitionistic propositional logic is an algebraic logic for the theory of Heyting algebras* Δ.

REMARK 6.2. *Note that an algebra with axioms* Δ *is written in a language with function symbols* $f_1, \ldots, f_n$ *and one monadic predicate* M. *The axioms of* Δ *may say something about* M. *We can add to the language an infinite number of monadic predicates* $P_1, P_2, \ldots$. *Call the new theory* Δ^M. *If we write no axioms about* $\{P_i\}$, *then we get that* Δ^M *is a conservative extension of* Δ. *Any algebra* $\mathfrak{A}$ *which is a model of* Δ *can also be a model of* Δ^M *and the* $\{P_i\}$, *we define them arbitrarily, and any model of* Δ^M *and* P_i *can be reduced to a model of* $\mathfrak{A}$ *alone by ignoring the* P_i.

We now define the translation of the algebraic logic $\mathfrak{L}$ into the theory of Δ augmented by P_i. Let h be a valuation, giving each atom p_i a variable v_i of the algebraic language. We translate the meta-statement:

The value $h(\varphi)$ of the formula φ in the algebra is x (denoted by $[\varphi]_x$).

Let us associate first with each atomic proposition p_i of $\mathfrak{L}$ a monadic predicate P_i of Δ.

https://doi.org/10.1017/S1755020308080246 Published online by Cambridge University Press

1. For atomic p_i, we have:
$$[p_i]_x = P_i(x),$$
where P_i is a monadic predicate associated with p_i.
2. $[c_i(A_1, \ldots, A_{m_i})]_x = \exists y_1, \ldots, y_{m_i}(\bigwedge_{i=1}^{m_i}[A_i]_{y_i} \wedge x = f_i(y_1, \ldots, y_{m_i}))$.

For each atom p_i of $\mathfrak{L}$, let Ψ_{p_i} be the following formula in classical logic:

$$\Psi_{p_i} = \exists x(P_i(x) \wedge \forall y(P_i(y) \rightarrow y = x)).$$

For any formula φ of $\mathfrak{L}$ with atoms $atoms(\varphi)$, let:

$$\tau(\varphi) = \bigwedge_{p_i \in atoms(\varphi)} \Psi_{p_i}.$$

For any formula φ of $\mathfrak{L}$, let Ψ_φ be:

$$\Psi_\varphi = \exists x[[\varphi]_x \wedge \forall y([\varphi]_y \rightarrow y = x)].$$

Note that this agrees with Ψ_p for atomic p since $[p]_x = P(x)$.

We can now prove by induction on φ the following lemmas.

LEMMA 6.3. *In classical logic,* $\Delta \models \tau(\varphi) \rightarrow \Psi_\varphi$.

LEMMA 6.4. *Let h be a valuation for $\mathfrak{L}$ giving value $h(p) \in \mathscr{M}$ for each atom p. Let $P(h(p))$ hold exactly for $h(p)$, for each atom p. Then, we have for any φ that $[\varphi]_{h(\varphi)}$ holds in the algebra.*

DEFINITION 6.5 ($\models_{\mathfrak{L}}$). *$\Delta \models_{\mathfrak{L}} \varphi$ if and only if $\forall \psi \in \Delta\ h(\psi) \in \mathscr{M} \backslash \mathscr{U}$ implies $h(\varphi) \in \mathscr{M} \backslash \mathscr{U}$.*

Note that in the case of Łukasiewicz's logic, this was the definition we did not want (see Remark 5.4). In the case of algebraic logic, we have no choice because we do not have specific knowledge about it. If the algebra have some well-known properties and they are problematic with this definition, then we can use them to define a more appropriate notion of consequence, as we did in the case of Łukasiewicz's logic.

DEFINITION 6.6. *A set Γ of well-formed formulae of $\mathfrak{L}$ is consistent if and only if for some algebra $\mathfrak{A}$ and some valuation h into $\mathfrak{A}$, we have that for every $\varphi \in \Gamma$, $h(\varphi) \in \mathfrak{D}$.*

Because we are speaking of a general algebra, we do not have a notion such as max, and we have to single out the designated values $\mathfrak{D}$. In case the algebra also has a set of undesignated values $\mathscr{U}$, the notion of consistency amounts to the existence of an assignment that gives a value in $\mathscr{M} \backslash \mathscr{U}$ (i.e., a non-undesignated value).

Let Γ be a theory of $\mathfrak{L}$. Let $\Gamma^\#$ be a theory of classical logic defined as follows:

$$\Gamma^\# = \{\tau(\varphi) \mid \varphi \in \Gamma\} \cup \{\exists y(D(y) \wedge [\varphi]_y) \mid \varphi \in \Gamma\},$$

where $D(y)$ holds if and only if $y \in \mathfrak{D}$.

LEMMA 6.7. *Γ is consistent in $\mathfrak{L}$ if and only if $\Delta^M \cup \Gamma^\#$ is classically consistent.*

6.2. Revision in algebraic logics.

DEFINITION 6.8. *Let $*$ be a revision process in classical logic, Γ a set of formulae of $\mathfrak{L}$, and φ a formula of $\mathfrak{L}$. Then, we can define a revision process in $\mathfrak{L}$ by $\Gamma * \varphi = \{A \in \Gamma \mid \Gamma^\# * (\Delta^M \wedge \tau(\varphi) \wedge \exists y(D(y) \wedge [\varphi]_y)) \vdash [A]_x$ for some $x \in \mathfrak{D}\}$.*

https://doi.org/10.1017/S1755020308080246 Published online by Cambridge University Press

EXAMPLE 6.9 (**Classical propositional logic**). *Consider the theory* Γ *with the three following formulae:*

1 $a \to b$
2 $a \to \neg b$
3 a.

Its translation into the theory of boolean algebra is as follows:

1^* $\exists y(D(y) \wedge \exists x_1, x_2(A(x_1) \wedge B(x_2) \wedge (y = \neg x_1 \vee x_2))$
2^* $\exists y(D(y) \wedge \exists x_1, x_2(A(x_1) \wedge B(x_2) \wedge (y = \neg x_1 \vee \neg x_2))$
3^* $\exists y(D(y) \wedge A(y))$.

Where $D(y)$ *says that* y *is designated (i.e.,* $y = 0$ *of the algebra if '*0*' is 'truth') and '*$=$*', '*$\wedge$*', '*$\vee$*', and '*$\neg$*' are boolean algebra operations.*

τ *says that* A *and* B *must hold for unique values, that is,* $\tau = \exists y(A(y) \wedge \forall u(A(u) \to u = y)) \wedge \exists y(B(y) \wedge \forall u(B(u) \to u = y))$.

Note that the set of formulae $\{\tau, 1^*, 2^*, 3^*\}$ *is consistent if and only if the set of formulae* $\{1, 2, 3\}$ *is consistent. If we cannot revise* τ *out, then we must revise* $\{1^*, 2^*, 3^*\}$.

6.3. Summary of the methodology. To provide a revision based on a translation into classical logic, we need:

1. Finite axioms for the algebra and for 0 (or for what it means to be a designated value).
2. Some axioms that constrain the uniqueness of the truth-value assignments to propositional atoms.
3. Finite number of atoms so that the above truth-value axioms are finite.

7. Revising Belnap's four-valued logic.

7.1. Belnap's four-valued logic. Standard familiar systems such as classical logic, modal logic, and intuitionistic logic have in common the principle that contradicting information entails any arbitrary sentence. This principle, known as *ex falsum quod libet*, is not, however, always appropriate to describe some deductive processes involving information coming from conflicting sources and alternative systems have been developed, including the logic of *entailment* (also known as system **E**) and the system of *relevant implication* (also known as system **R**). Deduction in these systems only holds when there is some 'connection' between the formulae (e.g., when they share some sentential variable).

Belnap (1977b) proposed a logic with four truth-values whose main motivation was the provision of a deduction mechanism that was less sensitive to inconsistencies. The truth-values of this logic are {None, F, T, Both} and represent three increasing levels of information one can gather about a proposition: From no information (None) one can learn that a proposition is true (T) or false (F), and from then any more information leads to a state where the proposition is both true and false (Both). Belnap (1977a) emphasised in his work its connection with the problem of 'how a computer should think'. A subsequent version of the work was then presented in Anderson *et al.* (1992).

Our objective in this section is to illustrate our approach by considering Belnap's four-valued logic, which we denote by $\mathbf{L}_B$. We start with a translation of the semantics of this logic into first-order logic via an appropriate axiomatisation capturing the behaviour of the connectives. This allows us to represent $\mathbf{L}_B$'s notion of entailment in terms of entailment

https://doi.org/10.1017/S1755020308080246 Published online by Cambridge University Press

in classical logic. The embedding into classical logic has two main advantages. First, it provides the basis for analysing belief revision operations for these types of logics, which is our main objective. However, it also allows for theorem proving in the four-valued logic using existing theorem provers for classical logic.

Syntax: The language $\mathscr{L}_B$ of Belnap's four-valued logic $\mathbf{L}_B$ is composed of a countable set of propositional letters $\mathscr{P}_B = \{p, q, r, \ldots\}$ and the connectives $\neg$, $\wedge$, and $\vee$. We use the symbol $\mathsf{wff}_{\mathscr{L}_B}$ to denote the set of all well-formed formulae of $\mathscr{L}_B$ (which are constructed in the standard way).

A finite set of well-formed formulae will usually be identified by the conjunction of all the formulae in the set (in any arrangement). We reserve lower case letters to represent the propositional symbols of $\mathscr{L}_B$ and lower case Greek letters for its well-formed formulae.

In proof theoretical terms, $\mathbf{L}_B$ is characterised by the finite axiomatisation given in the next section. The axiomatisation is known to be sound and complete with respect to the semantics presented later.

Given two formulae α and β, the expression $\alpha \rightarrow \beta$ will denote the fact that α *entails* β. You can regard the symbol $\rightarrow$ as a derivability relation between formulae or between sets of formulae and a formula. The expression $\alpha \leftrightarrow \beta$ will denote the fact that α entails β and vice versa (semantically, we say that α and β are equivalent).

7.2. Entailment in $\mathbf{L}_B$

DEFINITION 7.1 (**Axiomatisation**). *Let* $\alpha_1, \ldots, \alpha_n$, $\beta_1, \ldots, \beta_m$ *and* γ *be elements of* $\mathsf{wff}_{\mathscr{L}_B}$. *The following set of expressions provide a proof theory for* $\mathbf{L}_B$. *We refer to this set of axioms as* Ax_B*:*

1. $\neg\neg\alpha \leftrightarrow \alpha$.
2. $\neg(\alpha \wedge \beta) \leftrightarrow \neg\alpha \vee \neg\beta$.
3. $\neg(\alpha \vee \beta) \leftrightarrow \neg\alpha \wedge \neg\beta$.
4. $\alpha \vee \beta \leftrightarrow \beta \vee \alpha$.
5. $\alpha \vee (\beta \vee \gamma) \leftrightarrow (\alpha \vee \beta) \vee \gamma$.
6. $\alpha \vee (\beta \wedge \gamma) \leftrightarrow (\alpha \vee \beta) \wedge (\alpha \vee \gamma)$.
7. $\alpha \wedge \beta \leftrightarrow \beta \wedge \alpha$.
8. $\alpha \wedge (\beta \wedge \gamma) \leftrightarrow (\alpha \wedge \beta) \wedge \gamma$.
9. $\alpha \wedge (\beta \vee \gamma) \leftrightarrow (\alpha \wedge \beta) \vee (\alpha \wedge \gamma)$.
10. $\alpha_1 \wedge \cdots \wedge \alpha_n \rightarrow \beta_1 \vee \cdots \vee \beta_m$ *provided that* $\beta_j = \alpha_i$ *for some* i *and* j.
11. $\alpha \rightarrow \beta$ *and* $\beta \rightarrow \gamma$, *then* $\alpha \rightarrow \gamma$.
12. $\alpha \leftrightarrow \beta$ *and* $\beta \leftrightarrow \gamma$, *then* $\alpha \leftrightarrow \gamma$.
13. $\alpha \rightarrow \beta$ *if and only if* $\neg\beta \rightarrow \neg\alpha$.
14. $(\alpha \vee \beta) \rightarrow \gamma$ *if and only if* $\alpha \rightarrow \gamma$ *and* $\beta \rightarrow \gamma$.
15. $\alpha \rightarrow \beta$ *if and only if* $\beta \leftrightarrow (\alpha \vee \beta)$
16. $\alpha \rightarrow \beta$ *if and only if* $\alpha \leftrightarrow (\alpha \wedge \beta)$
17. $\alpha \rightarrow (\beta \wedge \gamma)$ *if and only if* $\alpha \rightarrow \beta$ *and* $\alpha \rightarrow \gamma$.

The first nine expressions correspond to standard classical properties of negation, disjunction, and conjunction (e.g., commutativity, associativity, De Morgan laws). We sometimes refer to them plus expression 10 as the axioms of $\mathbf{L}_B$. Expressions 10–13 capture reflexivity, transitivity, and contrapositive properties of $\rightarrow$, whereas Expressions

https://doi.org/10.1017/S1755020308080246 Published online by Cambridge University Press

14–17 correspond to standard classical rules for introduction and elimination of $\vee$ and $\wedge$, respectively. We sometimes refer to Expressions 11–17 as the rules of $\mathbf{L}_B$. Any expression of the form $\psi \rightarrow \varphi$ can be either an instantiation of one of the Axioms 1–10 or obtained using some of the Rules 11–17, together with some axiom instantiations. For any given expression $\psi \rightarrow \varphi$, we define the notion of *length of the proof* as the 'least number' of rule applications needed to show $\psi \rightarrow \varphi$.

The similarity between the aforementioned rules and the ones for classical logic shows that four-valued logics are indeed very close to standard classical logic. The basic classical rule that is missing in $\mathbf{L}_B$ and that makes it *paraconsistent* is the rule $(\alpha \wedge \neg\alpha) \rightarrow \beta$, often referred to as *ex falsum quod libel*. Within a classical framework, this rule allows the derivation of any arbitrary formula from contradictory assumptions.

Semantics: Let **4** be the set {None, T, F, Both} containing the four different truth-values, which an atomic sentence can have within a given 'state of information' as mentioned before. The intuitive meaning of these values for a proposition p is given as follows:

1. None: p's status is unknown.
2. T: p is known to be true only .
3. F: p is known to be false only .
4. Both: p is known to be both true and false.

The values in **4** form a lattice, called the *approximation lattice* and denoted by **A4** where the ordering relation $\sqsubseteq$ goes 'uphill' and respects the monotonicity property, in the sense that information about the truth-value of a formula always increases from None to Both (Figure 2).

The truth-values of complex formulae are defined according to the truth-tables shown in Figure 3. The tables are associated with another lattice, called the *logical lattice* and denoted by **L4** (Figure 2). In **L4**, logical conjunction and disjunction are associated with the meet and join operations, respectively.

A *set-up* s is a mapping of the atomic formulae into **4**. Using the truth-tables given in Figure 3, each set-up is extended to a mapping of *all formulae* into **4**, in the standard inductive way. Thus, a set-up plays the equivalent role of a valuation in classical propositional logic. What makes the semantics of $\mathbf{L}_B$ somewhat different from the classical logic one is that valuations that make formulae true (called models in classical logic) do not play any special role here. Models in classical logic can be used to define semantic entailment, whereas in $\mathbf{L}_B$, the notion of models does not exist and semantic entailment is expressed instead in terms of a partial ordering $\preceq$ associated with the logical lattice **L4**.

We use the symbol $\Rightarrow$ to denote semantic entailment distinguishing it from the proof theoretical one $\rightarrow$. The two notions are equivalent, as stated in (4) below, and the symbols $\rightarrow$ and $\Rightarrow$ can be used interchangeably (Belnap, 1977a).

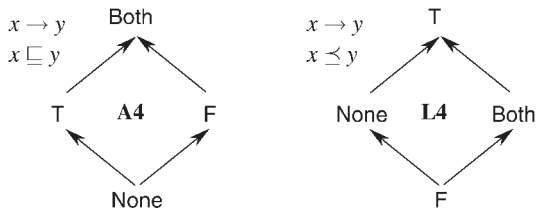


Fig. 2. The approximation (**A4**) and logical (**L4**) lattices.

https://doi.org/10.1017/S1755020308080246 Published online by Cambridge University Press

$\wedge$	None	F	T	Both
None	None	F	None	F
F	F	F	F	F
T	None	F	T	Both
Both	F	F	Both	Both

Truth-table for the connective $\wedge$

$\vee$	None	F	T	Both
None	None	None	T	T
F	None	F	T	Both
T	T	T	T	T
Both	T	Both	T	Both

Truth-table for the connective $\vee$

	None	F	T	Both
$\neg$	None	T	F	Both

Truth-table for the connective $\neg$

Fig. 3. Truth-tables for the connectives in $\mathbf{L}_B$.

$$\textbf{(Correspondence)} \quad \alpha \rightarrow \beta \text{ iff } \alpha \Rightarrow \beta. \tag{4}$$

DEFINITION 7.2. *Let α and β be formulae. We say that α semantically entails β, that is, $\alpha \Rightarrow \beta$, if for all set-ups s, $s(\alpha) \preceq s(\beta)$, where $\preceq$ is the partial ordering associated with* **L4**. *Analogously, a* non-empty *finite set of formulae* Γ semantically entails α, *if the conjunction of all formulae in* Γ *semantically entails* α.

Notice that if we restrict our attention to set-ups into $\{\mathsf{F}, \mathsf{T}\}$ only, we get the familiar classic logic notions (since $\mathsf{F} \preceq \mathsf{T}$ in **L4**).

Now, because this logic is paraconsistent, the translation of any set of its formulae into classical logic will always yield a consistent set (in classical logic) and a revision will never incur in removal of information (revisions would technically be equivalent to expansions).

We do not wish to delve into a long discussion into what would constitute a genuine need for 'revision' (in the non-conservative sense) for this kind of logics. For these considerations, the reader is referred to Restall & Slaney (1995), Fuhrmann (1991), and Priest (2001). However, we need to provide a notion to substitute classical consistency so that we can analyse the effects of a revision in this context. We call this *acceptability*. We say that a formula (or set of formulae) Δ of $\mathbf{L}_B$ is *acceptable*, if for all formulae α, it is not the case that Δ entails both α and $\neg\alpha$.

This somewhat ad hoc definition is sufficient for our purpose here. An acceptable theory Δ can tell us that a given formula α is true (T), false (F), or that its status is unknown (None), but if it says that it is both true and false (Both), then something in the knowledge about α went wrong and we may wish to *revise* Δ accordingly.

DEFINITION 7.3 (**Acceptability of theories of $\mathbf{L}_B$**). *A theory Δ is acceptable, if $\{\gamma \mid \Delta \rightarrow \gamma$ and $\Delta \rightarrow \neg\gamma\} = \emptyset$.*

It is worth emphasising that we do not have that if Δ is not acceptable, then it can prove everything. After all, this logic is paraconsistent, and hence, it will not trivialise in the presence of inconsistency. One may then ask why revise at all then? Our objective is to

https://doi.org/10.1017/S1755020308080246 Published online by Cambridge University Press

analyse what AGM revision would do in such a scenario. However, we feel that even though the logic will not trivialise in the presence of inconsistency, one may still decide to resolve it, either in one step (by defining acceptability as in Definition 7.3) or gradually (as suggested by Fuhrmann, 1991). The latter can be accomplished in our scenario by defining a weaker notion of acceptability.

We now introduce some terminology that are used later in this section.

DEFINITION 7.4. *Let α be a formula and s be a set-up.*

- *α is at least true under s if $s(\alpha) = \mathsf{T}$ or $s(\alpha) = \mathsf{Both}$.*
- *α is at least false under s if $s(\alpha) = \mathsf{F}$ or $s(\alpha) = \mathsf{Both}$.*
- *α is not true under s if $s(\alpha) = \mathsf{F}$ or $s(\alpha) = \mathsf{None}$.*
- *α is not false under s if $s(\alpha) = \mathsf{T}$ or $s(\alpha) = \mathsf{None}$.*

The notion of semantic entailment between a theory and a formula given in Definition 7.2 is defined in terms of $\preceq$ in **L4**. Since F is the least element in **L4**, all we need to watch out for is when the truth-value of the premises lies above F in the lattice. There are two cases to consider as defined below.

DEFINITION 7.5. *Let Γ be a set of formulae and α a formula. Γ semantically entails α if and only if for every set-up s,*

(i) If all the formulae in Γ are at least true under s, then α is at least true under s.
(ii) If all the formulae in Γ are not false under s, then α is not false under s.

If the truth-value of Γ is Both, then we expect the truth-value of α to be at least Both (i.e., at least true as in Part (i)). If the truth-value of Γ is T, then (i) ensures that the truth-value of α is at least true *and* (ii) that it is not false (i.e., that it is in fact T). On the other hand, if the truth-value of Γ is None, then we expect the truth-value of α to be at least None (i.e., not to be false as in Part (ii)).

Definition 7.5 will play an important role in the proof of soundness and completeness of our translation into first-order logic with respect to the semantics of $\mathbf{L}_B$.

7.3. Generalising the notion of entailment of $\mathbf{L}_B$ to the infinite case. In $\mathbf{L}_B$, the truth-value of a conjunction of formulae is defined inductively from the truth-values of its conjuncts. Alternatively, this value can be calculated by taking the meet of the truth-values of the conjuncts in **L4**. In that sense, finite sets of formulae and conjunctions behave in the same way. However, the notion of meet is more general because it can be applied to an infinite set of truth-values. This is an important point as we explain next.

Belief revision is an operation from a belief set and a formula to another belief set, and we know that a belief set is a set of formulae closed under the consequence relation. However, in doing our revision by translation, we do not have to start with the closure of a formula of $\mathscr{L}_B$ under $\mathbf{L}_B$'s consequence relation because our axiomatisation together with the closure in classical logic will take care of calculating that formula's consequences. This has, however, a side-effect. After we perform the revision in classical logic, we are left with a belief set (as required by (K*1)), which is an *infinite* set of formulae. When we translate back to the object language, we will consequently also get an infinite set formulae. To analyse the resulting operation under $\mathbf{L}_B$'s perspective, we need to ensure that $\mathbf{L}_B$'s notion of entailment ($\rightarrow$) is well-behaved for infinite sets. This is what we show in this section.

Since the evaluations of $\wedge$ and meet in **L4** coincide in $\mathbf{L}_B$, this will be our starting point.

https://doi.org/10.1017/S1755020308080246 Published online by Cambridge University Press

DEFINITION 7.6. *Let Δ be a (possibly infinite) set of formulae of $\mathscr{L}_B$, $\mathscr{S}$ the set of all set-ups, and s an arbitrary set-up. We define $val : \mathscr{S} \times \mathscr{P}(\mathsf{wff}_{\mathscr{L}_B}) \longmapsto \mathbf{4}$ as*[5]*:*

$$val(s, \Delta) = \sqcap\{s(\varphi) | \varphi \in \Delta\}.$$

We can now generalise $\mathbf{L}_B$'s notion of entailment as follows.

DEFINITION 7.7. *Let Δ be a (possibly infinite) set of formulae of $\mathscr{L}_B$ and φ a formula.*

$$\Delta \rightarrow \varphi \text{ iff } \forall s\ [val(s, \Delta) \preceq s(\varphi)],$$

where $\preceq$ is the partial order in **L4**.

Note that in the finite case, the aforementioned definition corresponds to the original notion of entailment of $\mathbf{L}_B$.

PROPOSITION 7.8. *Let Δ be a finite set of formulae and $\delta = \bigwedge_{\beta_i \in \Delta} \beta_i$. $val(s, \Delta) = s(\delta)$.*

Proof. The proof is done by induction on the size of Δ.
Base case: $|\Delta| = 1$. $\Delta = \{\delta'\}$, for some δ'. $val(s, \Delta) = \sqcap\{s(\delta')\}$. By the definition of $\sqcap$, $val(s, \Delta) = s(\delta') = s(\delta)$.

Inductive step: The proposition holds for any Γ such that $|\Gamma| \leq k$. Now, suppose that $|\Delta| = k + 1$. $\Delta = \Gamma \cup \{\beta\}$, for some Γ of cardinality k. Let $\gamma = \bigwedge_{\gamma_i \in \Gamma} \gamma_i$. It follows that $val(s, \Delta) = \sqcap\{s(\delta_i) \mid \delta_i \in \Delta\} = val(s, \Gamma) \sqcap s(\beta)$, by the distributivity property of $\sqcap$. By the induction hypothesis, $val(s, \Gamma) = s(\gamma)$. Thus, $val(s, \Delta) = s(\gamma) \sqcap s(\beta) = s(\gamma \wedge \beta) = s(\delta)$, by associativity and distributivity of $\wedge$. □

COROLLARY 7.9. *Let Δ be a finite set of formulae, φ a formula, and $\delta = \bigwedge_{\beta_i \in \Delta} \beta_i$. $val(s, \Delta) \preceq s(\varphi)$ iff $s(\delta) \preceq s(\varphi)$.*

The extended definition of entailment satisfies the finiteness property, as shown in Theorem 13. The proof of this property uses the following definitions and theorem.

DEFINITION 7.10 (**Ultra set-up**). *Let I be a non-empty set, D an ultrafilter over I, and s_i a set-up for $\mathbf{L}_B$, for each $i \in I$. $S = \Pi_D s_i$ is a function $S : \mathscr{P}_B \longmapsto \mathbf{4}$ defined as follows. For each propositional symbol $p \in \mathscr{P}_B$, $S(p) = u$ iff $\{i \in I \mid s_i(p) = u\} \in D$.*

The definition above is extended to complex formulae in the usual way, using $\mathbf{L}_B$'s truth-tables. For example, $S(\alpha \wedge \beta) = S(\alpha) \wedge S(\beta)$. It is useful to refer to the particular value of y assigned to a formula by the function S above. The following theorem shows that an ultra set-up as defined above is indeed a **4**-valuation as defined before.

THEOREM 7.11 (**The fundamental theorem**). *For each $\alpha \in \mathscr{L}_B$, $S(\alpha) = u$ if and only if $\{i \in I \mid s_i(\alpha) = u\} \in D$.*

Proof. The proof is done by induction on the complexity of α.
('Only if') The *base case* follows directly from Definition 7.10.
Assume that the theorem holds for any formula β with at most k connectives.

Inductive step: The proof is done by cases on the main connective of α.

($\neg$) α is of the form $\neg\beta$ for some β. Let $S(\neg\beta) = u$, for some u. By the definition of S, $S(\neg\beta) = \neg S(\beta) = \neg u_1$, for some $u_1 = S(\beta)$. Notice that by the inductive

[5] $\mathscr{P}(\mathsf{wff}_{\mathscr{L}_B})$ is the power set of the set of all well-formed formulae in $\mathscr{L}_B$.

https://doi.org/10.1017/S1755020308080246 Published online by Cambridge University Press

hypothesis, the set $N = \{n \mid s_n(\beta) = u_1\} \in D$. We show that the set $M = \{m \mid s_m(\neg\beta) = u\} = N$ and therefore $M \in D$. Let $k \in M$. Then, $s_k(\neg\beta) = \neg s_k(\beta) = \neg u_2$, for some u_2. By the definition of $\neg$, $\neg u_1 = \neg u_2$ implies that $u_1 = u_2$. Therefore, $k \in N$. Let $k \in N$. Then, $s_k(\beta) = u_1$. By definition of $\neg$, $s_k(\neg\beta) = \neg s_k(\beta) = \neg u_1 = u$. Hence, $k \in M$.

($\wedge$) α is of the form $\beta_1 \wedge \beta_2$, for some β_1 and β_2. Let $S(\alpha) = u$. By definition of S, $u = S(\beta_1) \wedge S(\beta_2)$. Suppose $S(\beta_1) = u_1$ and $S(\beta_2) = u_2$. By inductive hypothesis, $N_1 = \{n \mid s_n(\beta_1) = u_1\} \in D$ and $N_2 = \{n; \mid s_n(\beta_2) = u_2\} \in D$. Since D is a filter, the intersection $N_1 \cap N_2 \in D$. Hence, the set $K = \{k \mid s_k(\beta_1) = u_1 \text{ and } s_k(\beta_2) = u_2\} \in D$. We show that $M = \{m \mid s_m(\beta_1 \wedge \beta_2) = u_1 \wedge u_2\} \in D$. We prove this by showing that $M = K$. Let $m \in M$. Thus, $s_m(\beta_1 \wedge \beta_2) = u_1 \wedge u_2$. By definition of $\wedge$, $s_m(\beta_1 \wedge \beta_2) = s_m(\beta_1) \wedge s_m(\beta_2)$ and therefore $s_m(\beta_1) = u_1$ and $s_m(\beta_2) = u_2$. Then, $s_m \in N_1$ and $s_m \in N_2$ and hence $s_m \in K$. Assume that $k \in K$. By definition of K, $s_k(\beta_1) = u_1$ and $s_k(\beta_2) = u_2$. By definition of $\wedge$, $s_k(\beta_1 \wedge \beta_2) = u_1 \wedge u_2$. Hence, $k \in M$.

($\vee$) The proof is analogous to the previous case.

('If') The *base case* also follows directly from Definition 7.10.
Assume that the theorem holds for any formula β with at most k connectives.
Inductive step: The proof is done by cases on the main connective of α.

($\neg$) α is of the form $\neg\beta$ for some β. Let u be such that the set $N = \{n \mid s_n(\neg\beta) = u\} \in D$. By definition of $\neg$, the set $\{n \mid \neg s_n(\beta) = u\} \in D$. We show that the set $M = \{m \mid s_m(\beta) = \neg u\} \in D$. For suppose that $M \notin D$. By the property of the ultrafilter, the complement set $\{k \mid s_k(\beta) \neq \neg u\} \in D$. By definition of *neg*, the set $\{k \mid \neg s_k(\beta) \neq \neg\neg u\} \in D$.[6] Therefore, the set $\{k \mid s_k(\neg\beta) \neq u\} \in D$, which is a contradiction since its complement belongs to D.

($\wedge$) α is of the form $\beta_1 \wedge \beta_2$ for some β_1 and β_2. Let u be such that the set $N = \{n \mid s_n(\beta_1 \wedge \beta_2) = u\} \in D$. Let u_1 and u_2 be such that $M_1 = \{n \mid s_n(\beta_1) = u_1\} \in D$ and $M_2 = \{n \mid s_n(\beta_2) = u_2\} \in D$. By the inductive hypothesis $S(\beta_1) = u_1$ and $S(\beta_2) = u_2$, which implies that $S(\alpha) = u_1 \wedge u_2$. Furthermore, since D is a filter, the set $\{n \mid s_n(\beta_1) = u_1 \text{ and } s_n(\beta_2) = u_2\} \in D$. Therefore, the set $\{n \mid s_n(\beta_1 \wedge \beta_2) = u_1 \wedge u_2\} \in D$. Suppose that $u_1 \wedge u_2 \neq u$. Then, by the property of the ultrafilter, the complement $\{n \mid s_n(\beta_1 \wedge \beta_2) \neq u_1 \wedge u_2\} \notin D$ and hence $N \notin D$, which is a contradiction. Therefore, $S(\alpha) = u_1 \wedge u_2 = u$.

($\vee$) The proof is analogous to the previous case.

□

PROPOSITION 7.12. *Let s be a set-up and Δ an infinite set of formulae. There exists $\Delta_n \subset \Delta$, Δ_n finite, such that for all $\Delta_m \supseteq \Delta_n$, $s(\Delta_m) = s(\Delta_n)$.*

Proof. Suppose that for all Δ_n, there exists $\Delta_m \supseteq \Delta_n$ such that $s(\Delta_m) \neq s(\Delta_n)$. Since $s(\Delta_m) \leq s_m(\Delta_n)$, it follows that $s(\Delta_m) < s(\Delta_n)$. This will lead to a infinite (strict) descending sequence of values in **4**, which is a contradiction since **4** is finite. □

THEOREM 7.13 (**Finiteness**). *Let Δ be an infinite set of formulae and φ a formula in $\mathscr{L}_B$. If $\Delta \to \varphi$, then there exists $\Delta' \subset \Delta$ s.t. Δ' is finite and $\Delta' \to \varphi$.*

[6] Note that for any $x \in \mathbf{4}$, $\neg\neg x = x$.

https://doi.org/10.1017/S1755020308080246 Published online by Cambridge University Press

Proof. By contrapositive. We assume that:

$$\text{for all finite } \Delta' \subset \Delta,\ \Delta' \not\vdash \varphi, \tag{5}$$

and show that $\Delta \not\vdash \varphi$. By the definition of $\rightarrow$ and Equation (5) we get that $\forall \Delta_i \subset \Delta$, there exists s_i, such that $s_i(\Delta_i) \not\leq s_i(\varphi)$.

Let $\delta_0, \delta_1, \ldots$ be an enumeration of all formulae in Δ and I an index set constructed as follows: $I = \{\Delta_0, \Delta_1, \ldots, \Delta_n, \ldots\}$, where $\Delta_i = \{\delta_0, \ldots, \delta_i\}$. Let D be an ultrafilter over I, defined as $\{\hat{\Delta}_0, \hat{\Delta}_1, \ldots\}$, where $\hat{\Delta}_j = \{\Delta_i \mid i \geq j\}$. Consider the ultraproduct $S = \Pi_D s_i$. By Theorem 7.11, S behaves exactly as a normal set-up, and hence, for any formula α, $S(\alpha) = u_\alpha$ if and only if $\{m \mid s_m(\alpha) = u_\alpha\} \in D$.

We show that:

$$\text{for all } \Delta' \subset \Delta,\ \Delta' \text{ finite}, S(\Delta') \not\leq S(\varphi). \tag{6}$$

Suppose $\exists \Delta_n \subset \Delta$ s.t. $S(\Delta_n) \leq S(\varphi)$. Let $S(\Delta_n) = u_n$ for some u_n and $S(\varphi) = f$, for some f. By the definition of S, $N = \{j \mid s_j(\Delta_n) = u_n\} \in D$ and $F = \{k \mid s_k(\varphi) = f\} \in D$. By construction of D, the set $\hat{\Delta}_n = \{i \mid i \geq j\} \in D$ and therefore the set $M = N \cap \hat{\Delta}_n = \{m \mid m \geq j \text{ and } s_m(\Delta_n) = u_n\} \in D$. Notice that $\Delta_m \supset \Delta_n$ for all $m \in M$ and therefore $s_m(\Delta_m) \leq s_m(\Delta_n)$.

Consider the set $F \cap M$. Since D is a filter, $F \cap M = \{h \mid s_h(\varphi) = f \text{ and } s_h(\Delta_n) = u_n\} \in D$. In particular, by the construction of the ultraproduct,

(i) $s_h(\Delta_h) \leq s_h(\Delta_n) = u_n \leq f$
(ii) $s_h(\Delta_h) \not\leq s_h(\varphi) = f$,

which are in contradiction.

By Proposition 7.12, there exists $\Delta_n \in I$, such that for all $m \geq n$, $S(\Delta_m) = x$ for a certain $x \in \mathbf{4}$. Therefore, $S(\Delta) = x$, and by Equation (6), $x \not\leq f$. Hence, we have constructed a set-up S such that $S(\Delta) \not\leq S(\varphi)$, which contradicts the initial hypothesis. □

7.4. Translating $\mathbf{L}_B$ into classical logic. In this section, we describe a translation approach of $\mathbf{L}_B$ into first-order logic and show that it is sound and complete with respect to $\mathbf{L}_B$'s semantic notion of entailment. For this, we use a two-sorted first-order language $\mathscr{L}$ composed of the sort $\mathscr{F}$, called *B-formulae*, and the sort $\mathscr{V}$, called *truth-values*.

The set of constants of $\mathscr{F}$ is the set of propositional letters in $\mathscr{L}_B$, whereas terms of $\mathscr{F}$ are constructed using three main functions $\neg$, $\wedge$, and $\vee$, which correspond to the connectives in $\mathscr{L}_B$. The set of ground terms of $\mathscr{F}$ is thus equal to the set of well-formed formulae of $\mathscr{L}_B$. Similarly, the sort $\mathscr{V}$ is composed of the two constant symbols tt and ff, the basic constants from which $\mathbf{L}_B$'s four-valued semantics can be constructed. $\mathscr{L}$ also contains the two-sorted binary predicate $holds$, in which the first argument is a term from $\mathscr{F}$ and the second argument is a term from $\mathscr{V}$. $\mathscr{F}$ variables will be denoted by x, y, z, and so forth. First-order formulae are constructed in the usual way from the predicate $holds$ and the usual classical logic connectives.

Ground atomic formulae can be of two types $holds(\varphi, \mathsf{tt})$ and $holds(\varphi, \mathsf{ff})$ for any formula φ from $\mathscr{L}_B$. Atomic formulae of the first type mean that '$\mathsf{tt} \in v(\varphi)$', for some set-up s, which is equivalent to say that for some set-up s, the truth-value of φ in s is at least true. Analogously, the second type states that '$\mathsf{ff} \in s(\varphi)$', for some set-up s, which is equivalent to say that in s, the truth-value of φ is at least false. With these two types of atomic formulae, it is possible to express $\mathbf{L}_B$'s full semantics. For example, to say that a formula α has value Both in a set-up s, we can simply write $holds(\alpha, \mathsf{tt}) \wedge holds(\alpha, \mathsf{ff})$.

https://doi.org/10.1017/S1755020308080246 Published online by Cambridge University Press

The semantic behaviour of $\mathbf{L}_B$'s connectives is fully captured by the following first-order axiomatisation.

DEFINITION 7.14. *Given the two languages $\mathscr{L}_B$ and $\mathscr{L}$, the following six axioms constitute the first-order logic axiomatisation of $\mathbf{L}_B$'s semantics (in symbols $\mathscr{A}_B$):*

$$\forall x[holds(x, \mathsf{ff}) \leftrightarrow holds(\neg x, \mathsf{tt})] \quad \textit{(Ax 1)}$$
$$\forall x[holds(x, \mathsf{tt}) \leftrightarrow holds(\neg x, \mathsf{ff})] \quad \textit{(Ax 2)}$$
$$\forall x, y[holds(x \wedge y, \mathsf{tt}) \leftrightarrow (holds(x, \mathsf{tt}) \wedge holds(y, \mathsf{tt}))] \quad \textit{(Ax 3)}$$
$$\forall x, y[holds(x \wedge y, \mathsf{ff}) \leftrightarrow (holds(x, \mathsf{ff}) \vee holds(y, \mathsf{ff}))] \quad \textit{(Ax 4)}$$
$$\forall x, y[holds(x \vee y, \mathsf{tt}) \leftrightarrow (holds(x, \mathsf{tt}) \vee holds(y, \mathsf{tt}))] \quad \textit{(Ax 5)}$$
$$\forall x, y[holds(x \vee y, \mathsf{ff}) \leftrightarrow (holds(x, \mathsf{ff}) \wedge holds(y, \mathsf{ff}))] \quad \textit{(Ax 6).}$$

The *translation function* τ is a mapping from the set $\mathsf{wff}_{\mathscr{L}_B}$ to the set of ground atomic first-order formulae of the form $holds(\varphi, \mathsf{tt})$. For a given formula φ in $\mathscr{L}_B$, its translation in first-order logic φ^τ is the atomic formula $holds(\varphi, \mathsf{tt})$. A set of formulae in $\mathscr{L}_B$ is translated into first-order logic by translating each formula in the set. Translation of finite sets of formulae of $\mathscr{L}_B$ can be alternatively done by translating the conjunction of all the formulae in the set.

We now show that the aforementioned translation function together with the axiomatisation $\mathscr{A}_B$ is sound and complete with respect to $\mathbf{L}_B$'s notion of entailment.

THEOREM 7.15 (**Correspondence**). *Let ψ and φ be two formulae from $\mathsf{wff}_{\mathscr{L}_B}$.*

$$\psi \rightarrow \varphi \quad \textit{iff} \quad \mathscr{A}_B, holds(\psi, \mathsf{tt}) \vdash holds(\varphi, \mathsf{tt})$$
$$\textit{and} \quad \mathscr{A}_B, \neg holds(\psi, \mathsf{ff}) \vdash \neg holds(\varphi, \mathsf{ff}).$$

Proof. The 'if part' is given by Lemma 7.19, whereas the 'only if' part is given by Lemma 7.16. □

The aforementioned correspondence captures in first-order terms the notion of entailment given in Definition 7.5 whenever ψ is of the form $\alpha_1 \wedge \cdots \wedge \alpha_n$, with $\{\alpha_1, \ldots, \alpha_n\}$ being a set of formulae of $\mathscr{L}_B$. More specifically, for the first conjunct of the statement, the assumption $holds(\psi, \mathsf{tt})$ is equivalent, by Ax 3, to $holds(\alpha_1, \mathsf{tt}) \wedge \cdots \wedge holds(\alpha_n, \mathsf{tt})$, which can be read as 'all α_i's ($1 \leq i \leq n$) are at least true'. The consequence $holds(\varphi, \mathsf{tt})$ can also be read as 'φ is at least true'. Analogously, for the second half of the statement, the assumption $\neg holds(\psi, \mathsf{ff})$ is equivalent, by Ax 4, to $\neg holds(\alpha_1, \mathsf{ff}) \wedge \cdots \wedge \neg holds(\alpha_n, \mathsf{ff})$, where each $\neg holds(\alpha_i, \mathsf{ff})$ can be read as 'α_i is not false'.

LEMMA 7.16 (**Completeness**). *Let ψ and φ be two formulae of $\mathsf{wff}_{\mathscr{L}_B}$. If $\psi \rightarrow \varphi$, then $\mathscr{A}_B$, $holds(\psi, \mathsf{tt}) \vdash holds(\varphi, \mathsf{tt})$ and $\mathscr{A}_B, \neg holds(\psi, \mathsf{ff}) \vdash \neg holds(\varphi, \mathsf{ff})$.*

Proof. The proof is done by induction on the length n of the derivation $\psi \rightarrow \varphi$.
Base case: $n = 0$. Then, $\psi \rightarrow \varphi$ can only be an instantiation of one of the Axioms 1–10 given in Definition 7.1. The proof is therefore by cases on each of these axioms. Only some of the cases are shown here. The remaining ones can be proved using the same reasoning.

Case 1: $\psi \rightarrow \varphi$ is an instantiation of $\alpha_1 \wedge \cdots \wedge \alpha_h \rightarrow \beta_1 \vee \cdots \vee \beta_k$, for some h and k such that $\alpha_i = \beta_j$ for some i and j. We show in Figure 4 that $\mathscr{A}_B, holds(\alpha_1 \wedge \cdots \wedge \alpha_h, \mathsf{tt}) \vdash holds(\beta_1 \vee \cdots \vee \beta_k, \mathsf{tt})$ and in Figure 5 that $\mathscr{A}_B, \neg holds(\alpha_1 \wedge \cdots \wedge \alpha_h, \mathsf{ff}) \vdash \neg holds(\beta_1 \vee \ldots \vee \beta_k, \mathsf{ff})$.

Case 2: $\psi \rightarrow \varphi$ is an instantiation of $\alpha \vee (\beta \wedge \gamma) \rightarrow (\alpha \vee \beta) \wedge (\alpha \wedge \gamma)$. We show in Figure 7 that $\mathscr{A}_B, holds(\alpha \vee (\beta \wedge \gamma), \mathsf{tt}) \rightarrow holds((\alpha \vee \beta) \wedge (\alpha \wedge \gamma), \mathsf{tt})$ and in Figure 6

https://doi.org/10.1017/S1755020308080246 Published online by Cambridge University Press

$$\frac{\mathscr{A}_B, holds(\alpha_1 \wedge \ldots \wedge \alpha_h, \mathtt{tt})}{holds(\alpha_1, \mathtt{tt}) \wedge \ldots \wedge holds(\alpha_h, \mathtt{tt})}\ (\text{Ax } 3)$$

$$\frac{holds(\alpha_1, \mathtt{tt}) \wedge \ldots \wedge holds(\alpha_h, \mathtt{tt})}{holds(\alpha_i, \mathtt{tt})}\ (\mathscr{E}\wedge)$$

$$\frac{holds(\alpha_i, \mathtt{tt})}{holds(\beta_j, \mathtt{tt})}\ (\text{equiv. rewriting})$$

$$\frac{holds(\beta_j, \mathtt{tt})}{holds(\beta_1, \mathtt{tt}) \vee \ldots \vee holds(\beta_k, \mathtt{tt})}\ (\mathscr{I}\vee)$$

$$\frac{holds(\beta_1, \mathtt{tt}) \vee \ldots \vee holds(\beta_k, \mathtt{tt})}{holds(\beta_1 \vee \ldots \vee \beta_k, \mathtt{tt})}\ (\text{Ax } 5)$$

Fig. 4. First-order proof of $\mathbf{L}_B$'s Axiom 10.

$$\frac{\mathscr{A}_B, \neg holds(\alpha_1 \wedge \ldots \wedge \alpha_h, \mathtt{ff})}{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{holds(\beta_1 \vee \ldots \vee \beta_k, \mathtt{ff})}{holds(\beta_1, \mathtt{ff}) \wedge \ldots \wedge holds(\beta_k, \mathtt{ff})}\ (\text{Ax } 6)}{holds(\beta_j, \mathtt{ff})}\ (\mathscr{E}\wedge)}{holds(\alpha_i, \mathtt{ff})}\ (\text{equiv. rewriting})}{holds(\alpha_1, \mathtt{ff}) \vee \ldots \vee holds(\alpha_h, \mathtt{ff})}\ (\mathscr{I}\vee)}{holds(\alpha_1 \wedge \ldots \wedge \alpha_h, \mathtt{ff})}\ (\text{Ax } 4)}{\bot}\ (\mathscr{I}\neg)}{\neg holds(\beta_1 \vee \ldots \vee \beta_k, \mathtt{ff})}$$

Fig. 5. First-order proof of $\mathscr{L}_B$'s Axiom 10.

$$\frac{\mathscr{A}_B, \neg holds(\alpha \vee (\beta \wedge \gamma), \mathtt{ff})}{\neg holds(\alpha, \mathtt{ff}) \vee \neg holds(\beta \wedge \gamma, \mathtt{ff})}\ (\text{Ax } 6)$$

$$\frac{\neg holds(\alpha, \mathtt{ff}) \vee \neg holds(\beta \wedge \gamma, \mathtt{ff})}{\neg holds(\alpha, \mathtt{ff}) \vee (\neg holds(\beta, \mathtt{ff}) \wedge \neg holds(\gamma, \mathtt{ff}))}\ (\text{Ax } 4)$$

$$\frac{\neg holds(\alpha, \mathtt{ff}) \vee (\neg holds(\beta, \mathtt{ff}) \wedge \neg holds(\gamma, \mathtt{ff}))}{(\neg holds(\alpha, \mathtt{ff}) \vee \neg holds(\beta, \mathtt{ff})) \wedge (\neg holds(\alpha, \mathtt{ff}) \vee \neg holds(\gamma, \mathtt{ff}))}\ (\text{De Morgan Law})$$

$$\frac{(\neg holds(\alpha, \mathtt{ff}) \vee \neg holds(\beta, \mathtt{ff})) \wedge (\neg holds(\alpha, \mathtt{ff}) \vee \neg holds(\gamma, \mathtt{ff}))}{\neg holds(\alpha \vee \beta, \mathtt{ff}) \wedge \neg holds(\alpha \vee \gamma, \mathtt{ff})}\ (\text{Ax } 6)$$

$$\frac{\neg holds(\alpha \vee \beta, \mathtt{ff}) \wedge \neg holds(\alpha \vee \gamma, \mathtt{ff})}{\neg holds((\alpha \vee \beta) \wedge (\alpha \vee \gamma), \mathtt{ff})}\ (\text{Ax } 4)$$

Fig. 6. First-order proof of left-to-right part of $\mathbf{L}_B$'s Axiom 6.

that $\mathscr{A}_B, \neg holds(\alpha \vee (\beta \wedge \gamma), \mathtt{ff}) \rightarrow \neg holds((\alpha \vee \beta) \wedge (\alpha \wedge \gamma), \mathtt{ff})$. Similar argument is applied in the case where $\psi \rightarrow \varphi$ is an instantiation of $(\alpha \vee \beta) \wedge (\alpha \wedge \gamma) \rightarrow \alpha \vee (\beta \wedge \gamma)$.

Inductive step: We assume that there exists a first part of a derivation proving an expression of the form $\alpha \rightarrow \beta$ with $n - 1$ applications of $\mathbf{L}_B$'s rules and that the n-th application of

https://doi.org/10.1017/S1755020308080246 Published online by Cambridge University Press

$$\frac{\mathscr{A}_B, holds(\alpha \vee (\beta \wedge \gamma), \mathsf{tt})}{holds(\alpha, \mathsf{tt}) \vee holds(\beta \wedge \gamma, \mathsf{tt})} \text{ (Ax 5)}$$

$$\frac{holds(\alpha, \mathsf{tt}) \vee holds(\beta \wedge \gamma, \mathsf{tt})}{holds(\alpha, \mathsf{tt}) \vee (holds(\beta, \mathsf{tt}) \wedge holds(\gamma, \mathsf{tt}))} \text{ (Ax 3)}$$

$$\frac{holds(\alpha, \mathsf{tt}) \vee (holds(\beta, \mathsf{tt}) \wedge holds(\gamma, \mathsf{tt}))}{(holds(\alpha, \mathsf{tt}) \vee holds(\beta, \mathsf{tt})) \wedge (holds(\alpha, \mathsf{tt}) \vee holds(\gamma, \mathsf{tt}))} \text{ (De Morgan Law)}$$

$$\frac{(holds(\alpha, \mathsf{tt}) \vee holds(\beta, \mathsf{tt})) \wedge (holds(\alpha, \mathsf{tt}) \vee holds(\gamma, \mathsf{tt}))}{holds(\alpha \vee \beta, \mathsf{tt}) \wedge holds(\alpha \vee \gamma, \mathsf{tt})} \text{ (Ax 5)}$$

$$\frac{holds(\alpha \vee \beta, \mathsf{tt}) \wedge holds(\alpha \vee \gamma, \mathsf{tt})}{holds((\alpha \vee \beta) \wedge (\alpha \vee \gamma), \mathsf{tt})} \text{ (Ax 3)}$$

Fig. 7. First-order proof of left-to-right part of $\mathbf{L}_B$'s Axiom 6.

rule of $\mathbf{L}_B$ gives us the expression $\psi \rightarrow \varphi$. We reason by cases on each rule that could have been applied on the n-th step.

Case 1: We assume that last rule application is the 'if part' of $\mathbf{L}_B$'s Rule 13 in Definition 7.1. Therefore, we have that there exists a proof of $\neg\varphi \rightarrow \neg\psi$, with $n - 1$ rule applications. So by inductive hypothesis, we can say that $\mathscr{A}_B, holds(\neg\varphi, \mathsf{tt}) \vdash holds(\neg\psi, \mathsf{tt})$ and that $\mathscr{A}_B, \neg holds(\neg\varphi, \mathsf{ff}) \vdash \neg holds(\neg\psi, \mathsf{ff})$. We want then to show that:

$$\mathscr{A}_B, holds(\psi, \mathsf{tt}) \vdash holds(\varphi, \mathsf{tt}),$$

and that

$$\mathscr{A}_B, \neg holds(\psi, \mathsf{ff}) \vdash \neg holds(\varphi, \mathsf{ff}).$$

From the inductive hypothesis $\mathscr{A}_B, \neg holds(\neg\varphi, \mathsf{ff}) \vdash \neg holds(\neg\psi, \mathsf{ff})$, we get, by contrapositive of classical logic, that:

$$\mathscr{A}_B, holds(\neg\psi, \mathsf{ff}) \vdash holds(\neg\varphi, \mathsf{ff})$$

Hence, using $\mathbf{L}_B$'s Axiom 2, we get $\mathscr{A}_B, holds(\psi, \mathsf{tt}) \vdash holds(\varphi, \mathsf{tt})$. To show that $\mathscr{A}_B, \neg holds(\psi, \mathsf{ff}) \vdash \neg holds(\varphi, \mathsf{ff})$, we consider the second part of the inductive hypothesis. $\mathscr{A}_B, holds(\neg\varphi, \mathsf{tt}) \vdash holds(\neg\psi, \mathsf{tt})$ gives, by contrapositive of classical logic, that:

$$\mathscr{A}_B, \neg holds(\neg\psi, \mathsf{tt}) \vdash \neg holds(\neg\varphi, \mathsf{tt}).$$

Hence, by $\mathbf{L}_B$'s Axiom 1, $\mathscr{A}_B, \neg holds(\psi, \mathsf{ff}) \vdash \neg holds(\varphi, \mathsf{ff})$. The case for the 'only if part' of $\mathbf{L}_B$'s Rule 13 follows the same reasoning.

Case 2: We assume that last rule application is the 'if part' $\mathbf{L}_B$'s Rule 15 in Definition 7.1. Therefore, we have there exists a proof of $\alpha \rightarrow \beta$ with $n - 1$ rule applications, where ψ is equal to α and φ is equal to $\alpha \vee \beta$. So by inductive hypothesis, $\mathscr{A}_B, holds(\alpha, \mathsf{tt}) \vdash holds(\beta, \mathsf{tt})$ and that $\mathscr{A}_B, \neg holds(\alpha, \mathsf{ff}) \vdash \neg holds(\beta, \mathsf{ff})$. We want to show that:

1. $\mathscr{A}_B, holds(\beta, \mathsf{tt}) \vdash holds(\alpha \vee \beta, \mathsf{tt})$ and
   $\mathscr{A}_B, holds(\alpha \vee \beta, \mathsf{tt}) \vdash holds(\beta, \mathsf{tt})$
2. $\mathscr{A}_B, \neg holds(\beta, \mathsf{ff}) \vdash \neg holds(\alpha \vee \beta, \mathsf{ff})$ and
   $\mathscr{A}_B, \neg holds(\alpha \vee \beta, \mathsf{ff}) \vdash \neg holds(\beta, \mathsf{ff})$.

The first part of (1) is quite straightforward. We show the second part. Assume $\mathscr{A}_B, holds(\alpha \vee \beta, \mathsf{tt})$. By Ax 5 and reflexivity of classical logic, $\mathscr{A}_B, holds(\alpha \vee \beta, \mathsf{tt}) \vdash \mathscr{A}_B, holds(\alpha, \mathsf{tt}) \vee holds(\beta, \mathsf{tt})$. By the inductive hypothesis, we have that $\mathscr{A}_B, holds(\alpha, \mathsf{tt}) \vdash holds(\beta, \mathsf{tt})$ and by reflexivity of classical logic:

$$\mathscr{A}_B, holds(\beta, \mathsf{tt}) \vdash holds(\beta, \mathsf{tt}).$$

https://doi.org/10.1017/S1755020308080246 Published online by Cambridge University Press

Therefore, using classical $\vee$-introduction rule,

$$\mathscr{A}_B, holds(\alpha, \mathsf{tt}) \vee holds(\beta, \mathsf{tt}) \vdash holds(\beta, \mathsf{tt}).$$

Hence, $\mathscr{A}_B, holds(\alpha \vee \beta, \mathsf{tt}) \vdash holds(\beta, \mathsf{tt})$. The proof for Part (2) follows the same argument.
All the other cases can be easily proved using appropriate properties and rules of classical logic and, if necessary, the axioms of $\mathbf{L}_B$. □

The soundness of the classical translation is based on the idea that for any given set-up of $\mathbf{L}_B$, it is always possible to construct a classical interpretation I, which satisfies the axioms $\mathscr{A}_B$ and preserves $\mathbf{L}_B$'s properties of semantic entailment. We first show how this classical logic interpretation can be constructed, then we show some of its properties.

DEFINITION 7.17. *Let s be a set-up of $\mathbf{L}_B$. A classical logic interpretation associated with s, and denoted by $\mathscr{I}_s$, is a function defined as follows:*

- $\mathscr{I}_s(\mathsf{tt}) = \mathsf{tt}$ *and* $\mathscr{I}_s(\mathsf{ff}) = \mathsf{ff}$.

Also, for each ground term α of sort $\mathscr{F}$:

- $\mathscr{I}_s(\alpha) = \alpha$, *for each ground term α of sort $\mathscr{F}$.*
- $\mathscr{I}_s(holds) = \{\langle \alpha, \mathsf{tt}\rangle \mid \mathsf{tt} \in v(\alpha)\} \cup \{\langle \alpha, \mathsf{ff}\rangle \mid \mathsf{ff} \in v(\alpha))\}$.

It is easy to show, by definition of $\mathscr{I}_s$, that the following properties hold for any formula α in $\mathscr{L}_B$ and any set-up s.

- $s(\alpha) = \mathsf{T}$ if and only if $\mathscr{I}_s \models holds(\alpha, \mathsf{tt}) \wedge \neg holds(\alpha, \mathsf{ff})$.
- $s(\alpha) = \mathsf{F}$ if and only if $\mathscr{I}_s \models holds(\alpha, \mathsf{ff}) \wedge \neg holds(\alpha, \mathsf{tt})$.
- $s(\alpha) = \mathsf{Both}$ if and only if $\mathscr{I}_s \models holds(\alpha, \mathsf{tt}) \wedge holds(\alpha, \mathsf{ff})$.
- $s(\alpha) = \mathsf{None}$ if and only if $\mathscr{I}_s \models \neg holds(\alpha, \mathsf{tt}) \wedge \neg holds(\alpha, \mathsf{ff})$.

The following proposition shows that a classical interpretation $\mathscr{I}_s$ associated to a given set-up s is a model of the first-order axioms $\mathscr{A}_B$.

PROPOSITION 7.18. *Let s be a set-up and $\mathscr{I}_s$ be its associated classical logic interpretation. Then, $\mathscr{I}_s$ is a model of the classical logic axiomatisation $\mathscr{A}_B$.*

Proof. The proof is done by cases of each axiom of $\mathscr{A}_B$.

Case 1 (Ax 1): We want to show that $\mathscr{I}_s \models \forall x[holds(x, \mathsf{ff}) \leftrightarrow holds(\neg x, \mathsf{tt})]$. We reason by contradiction. We assume that for some x, $\mathscr{I}_s \models holds(x, \mathsf{ff})$ and $\mathscr{I}_s \not\models holds(\neg x, \mathsf{tt})$. By definition of $\mathscr{I}_s$, $\mathsf{ff} \in v(x)$, which implies by the $\neg$ truth-table that $\mathsf{tt} \in v(\neg x)$. Hence, $\mathscr{I}_s \models holds(\neg x, \mathsf{tt})$, which contradicts the hypothesis. Similarly for the other case, that is, $\mathscr{I}_s \not\models holds(x, \mathsf{ff})$ and $\mathscr{I}_s \models holds(\neg x, \mathsf{tt})$.

Case 3 (Ax 3): We want to show that $\mathscr{I}_s \models \forall x, y[holds(x \wedge y, \mathsf{tt}) \leftrightarrow (holds(x, \mathsf{tt}) \wedge holds(y, \mathsf{tt}))]$. We reason by contradiction. Assume that for some x, $\mathscr{I}_s \models holds(x \wedge y, \mathsf{tt})$ and $\mathscr{I}_s \not\models holds(x, \mathsf{tt})$ or $\mathscr{I}_s \not\models holds(y, \mathsf{tt})$. By definition of $\mathscr{I}_s$, $\mathsf{tt} \in v(x \wedge y)$, which implies by the $\wedge$ truth-table that $\mathsf{tt} \in v(x)$ and $\mathsf{tt} \in v(y)$. Therefore, $\mathscr{I}_s \models holds(x, \mathsf{tt})$ and $\mathscr{I}_s \models holds(y, \mathsf{tt})$, which is in contradiction with the initial hypothesis. The second case, that is, assume that for some x, $\mathscr{I}_s \not\models holds(x \wedge y, \mathsf{tt})$, and $\mathscr{I}_s \models holds(x, \mathsf{tt})$ and $\mathscr{I}_s \models holds(y, \mathsf{tt})$, can be proved following the same argument.

Case 5 (Ax 5): We want to show that $\mathscr{I}_s \models \forall x, y[holds(x \vee y, \mathsf{tt}) \leftrightarrow (holds(x, \mathsf{tt}) \vee holds(y, \mathsf{tt}))]$. We reason by contradiction. Assume that for some x, $\mathscr{I}_s \models holds(x \vee y, \mathsf{tt})$,

https://doi.org/10.1017/S1755020308080246 Published online by Cambridge University Press

and $\mathscr{I}_s \not\models holds(x, \mathsf{tt})$ and $\mathscr{I}_s \not\models holds(y, \mathsf{tt})$. By definition of $\mathscr{I}_s$, $\mathsf{tt} \in v(x \vee y)$, which implies by the $\vee$ truth-table that $\mathsf{tt} \in v(x)$ or $\mathsf{tt} \in v(y)$. Therefore, $\mathscr{I}_s \models holds(x, \mathsf{tt})$ or $\mathscr{I}_s \models holds(y, \mathsf{tt})$, which is in contradiction with the initial hypothesis. The second case, that is, assume that for some x, $\mathscr{I}_s \not\models holds(x \vee y, \mathsf{tt})$, and $\mathscr{I}_s \models holds(x, \mathsf{tt})$ or $\mathscr{I}_s \models holds(y, \mathsf{tt})$, can be proved following the same argument.

Axioms 2,4, and 6 are proved in an analogous way of the proofs of Axioms, 1,3, and 5, respectively. □

LEMMA 7.19 (**Soundness**). *Let ψ and φ be two Belnap formulae. If $\mathscr{A}_B, holds(\psi, \mathsf{tt}) \vdash holds(\varphi, \mathsf{tt})$ and $\mathscr{A}_B, \neg holds(\psi, \mathsf{ff}) \vdash \neg holds(\varphi, \mathsf{ff})$, then $\psi \rightarrow \varphi$.*

Proof. We prove the contrapositive statement. We assume that $\psi \not\rightarrow \varphi$ and we want to show that either $\mathscr{A}_B, holds(\psi, \mathsf{tt}) \not\vdash holds(\varphi, \mathsf{tt})$ or $\mathscr{A}_B, \neg holds(\psi, \mathsf{ff}) \not\vdash \neg holds(\varphi, \mathsf{ff})$. The hypothesis $\psi \not\rightarrow \varphi$ implies different possibilities of truth-values for ψ and φ according to the ordering relation $\preceq$ over the logical lattice **L4**. We consider these cases individually. $\psi \not\rightarrow \varphi$ implies that for some set-up s, $s(\psi) \not\preceq s(\varphi)$.

Case 1: $s(\psi) = \mathsf{T}$ and $s(\varphi) = \mathsf{Both}$. From s, we can construct the associated classical interpretation $\mathscr{I}_s$. By definition, $\mathscr{I}_s \models holds(\psi, \mathsf{tt})$ and $\mathscr{I}_s \models \neg holds(\psi, \mathsf{ff})$. But $\mathscr{I}_s \not\models \neg holds(\varphi, \mathsf{ff})$.

Case 2: $s(\psi) = \mathsf{T}$ and $s(\varphi) = \mathsf{None}$. Then, $\mathsf{tt} \notin s(\varphi)$. From s, we can construct the associated classical interpretation $\mathscr{I}_s$. By definition, $\mathscr{I}_s \models holds(\psi, \mathsf{tt})$ and $\mathscr{I}_s \models \neg holds(\psi, \mathsf{ff})$. But $\mathscr{I}_s \not\models holds(\varphi, \mathsf{tt})$.

Case 3: $s(\psi) = \mathsf{T}$ and $s(\varphi) = \mathsf{F}$. Then, $\mathsf{tt} \notin s(\varphi)$. From s, we can construct the associated classical interpretation $\mathscr{I}_s$. By definition, $\mathscr{I}_s \models holds(\psi, \mathsf{tt})$ and $\mathscr{I}_s \models \neg holds(\psi, \mathsf{ff})$. But $\mathscr{I}_s \not\models holds(\varphi, \mathsf{tt})$.

Case 4: $s(\psi) = \mathsf{None}$ and $s(\varphi) = \mathsf{F}$. Then, $\mathsf{tt} \notin s(\psi)$, $\mathsf{ff} \notin s(\psi)$, and $\mathsf{tt} \notin s(\varphi)$. From s, we can construct the associated classical interpretation $\mathscr{I}_s$. By definition, $\mathscr{I}_s \models \neg holds(\psi, \mathsf{ff})$ but $\mathscr{I}_s \not\models \neg holds(\varphi, \mathsf{ff})$.

Case 5: $s(\psi) = \mathsf{Both}$ and $s(\varphi) = \mathsf{F}$. Then, $\mathsf{tt} \in s(\psi)$, $\mathsf{ff} \in s(\psi)$, and $\mathsf{tt} \notin s(\varphi)$. From s, we can construct the associated classical interpretation $\mathscr{I}_s$. By definition, $\mathscr{I}_s \models holds(\psi, \mathsf{tt})$, but but $\mathscr{I}_s \not\models holds(\varphi, \mathsf{tt})$. □

7.5. Revision in Belnap's four-valued logic. We are now in a position to analyse in detail how our approach behaves with respect to Belnap's four-valued logic $\mathbf{L}_B$. Part of our motivation here is to analyse the role of inconsistency in the revision process. We have already mentioned that because of postulate (K*5), the result of an AGM revision is always consistent provided that the revising formula is not itself contradictory. More importantly, if the belief set being revised is consistent with the new formula, then nothing is removed in the revision process. In other words, the revision is equivalent to an expansion. This may cause problems if the notion of consistency of the object language differs significantly from that of classical logic's. In such cases, we need to seek for alternatives to *trigger* the revision process in the translated belief set.

Our view is that one does not need to be limited to a particular notion of consistency, as long as some notion of *acceptability* of theories in the object language can be defined via classical logic axioms.

https://doi.org/10.1017/S1755020308080246 Published online by Cambridge University Press

The idea now is to map the notion of acceptability given in Definition 7.3 to classical logic in such a way that the translation of theories in $\mathbf{L}_B$ that are not acceptable yields *inconsistent* classical logic theories, in which case a 'true' revision will be performed. We also need to ensure that the acceptability notion is included in the revising argument of the operation, so that by (K*2), the result of the revision will be itself *acceptable*, as long as the input is itself acceptable. Our notion of acceptability can be translated to classical logic in the following way.

DEFINITION 7.20. *The acceptability axiom for* $\mathbf{L}_B$ *is given by the formula Acc below.*

$$Acc = \forall\alpha[(holds(\alpha, \mathsf{tt}) \wedge holds(\alpha, \mathsf{ff})) \rightarrow \bot].$$

The co-existence of contradictory information in the object theory will not be tolerated by the revision process, and the new belief set will contain only formulae that are assigned *at most* one of the truth-values tt and ff (effectively filtering out formulae with truth-value Both). To simplify notation, we use the following convention:

$$holds(\alpha, \mathsf{tt}) = \alpha^+ \; holds(\neg\alpha, \mathsf{tt}) = \overline{\alpha}^+$$
$$holds(\alpha, \mathsf{ff}) = \alpha^- \; holds(\neg\alpha, \mathsf{ff}) = \overline{\alpha}^-.$$

Furthermore, we use $\neg\alpha^+$ for $\neg holds(\alpha, \mathsf{tt})$, and so forth. Now, belief revision in $\mathbf{L}_B$ logic is defined as follows.

DEFINITION 7.21 (**Belief revision in** $\mathbf{L}_B$). *Let* ψ *and* φ *be elements of* $\mathsf{wff}_{\mathscr{L}_B}$ *and* ψ^+ *(*ψ^-*) and* φ^+ *(*φ^-*) their translation to classical logic as described above, respectively. Let* $\mathscr{A}_B$ *be as in Definition 7.14 and Acc as in Definition 7.20. The revision of* ψ *by* φ*, in symbols* $\psi *_b \varphi$*, is defined as:*

$$\psi *_b \varphi = \{\gamma \mid \psi^* {}_a(\varphi^+ \wedge \mathscr{A}_B \wedge Acc) \vdash \gamma^+ \text{ and } \neg\psi^- *_a(\neg\varphi^- \wedge \mathscr{A}_B \wedge Acc) \vdash \neg\gamma^-.\}$$

Motivation. Intuitively speaking, the positive *holds* checks the relevant part of one-half of the lattice **L4**, that is, Both and T, whereas the negative *holds* checks the other half, that is, None and F. F does not need to be checked because it is the minimum in **L4**.

Notice that without *Acc*, the revision of ψ by φ is simply $\psi \wedge \varphi$. The correspondence axiom tell us that $\psi \wedge \varphi \rightarrow \gamma$ if and only if $\mathscr{A}_B \wedge (\psi \wedge \varphi)^+ \vdash \gamma^+$ and $\mathscr{A}_B \wedge \neg(\psi \wedge \varphi)^- \vdash \neg\gamma^-$. Since $\varphi^+ \wedge \mathscr{A}_B$ is consistent with ψ^+, $\psi^* {}_a(\varphi^+ \wedge \mathscr{A}_B)$ is just $\psi^+ \wedge \varphi^+ \wedge \mathscr{A}_B$ (the same for the negative part). The revision then gives us all $\gamma's$ such that $\psi^+ \wedge \varphi^+ \wedge \mathscr{A}_B \vdash \gamma^+$ and $\neg\psi^- \wedge \neg\varphi^- \wedge \mathscr{A}_B \vdash \neg\gamma^-$. Notice that in the presence of $\mathscr{A}_B$, $\psi^+ \wedge \varphi^+ \longleftrightarrow (\psi \wedge \varphi)^+$ and $\neg\psi^- \wedge \neg\varphi^- \longleftrightarrow \neg(\psi \wedge \varphi)^-$ (from Ax 3 and Ax 4, respectively). Therefore, we get all $\gamma's$ such that $\mathscr{A}_B \wedge (\psi \wedge \varphi)^+ \vdash \gamma^+$ and $\mathscr{A}_B \wedge \neg(\psi \wedge \varphi)^- \vdash \neg\gamma^-$, which is $\{\gamma \mid \psi \wedge \varphi \rightarrow \gamma\}$. This is how it should be because $\mathbf{L}_B$ is paraconsistent.

7.6. Analysis of the revision operation obtained via the translation. In this section, we analyse properties of the revision operator presented in Definition 7.21 with respect to the original AGM postulates.

(K*1) $\psi *_b \varphi$ is a belief set.

In our terms, this means that the set of formulae of $\mathscr{L}_B$ that we receive back from the revision process should be closed under $\rightarrow$. In other words, $\gamma \in \psi *_b \varphi$ if and only if $\psi *_b \varphi \rightarrow \gamma$. (Only if) If $\gamma \in \psi *_b \varphi$, it follows by reflexivity of $\rightarrow$ that $\psi *_b \varphi \rightarrow \gamma$. (If) Suppose that $\psi *_b \varphi \rightarrow \gamma$, we need to show that $\gamma \in \psi *_b \varphi$. By the correspondence

https://doi.org/10.1017/S1755020308080246 Published online by Cambridge University Press

theorem, if $\psi *_b \varphi \rightarrow \gamma$, then $\{\mathscr{A}_B\} \cup \{\alpha^+ \mid \psi^+ *_a (\varphi^+ \wedge \mathscr{A}_B \wedge Acc) \vdash \alpha^+\} \vdash \gamma^+$ and $\{\mathscr{A}_B\} \cup \{\neg\alpha^- \mid \neg\psi^- *_a (\neg\varphi^- \wedge \mathscr{A}_B \wedge Acc) \vdash \neg\alpha^-\} \vdash \neg\gamma^-$.

Notice that $\{\mathscr{A}_B\} \cup \{\alpha^+ \mid \psi^+ *_a (\varphi^+ \wedge \mathscr{A}_B \wedge Acc) \vdash \alpha^+\} \subseteq \psi^+ *_a (\varphi^+ \wedge \mathscr{A}_B \wedge Acc)$ since by (K*2), $\mathscr{A}_B \in \psi^+ *_a (\varphi^+ \wedge \mathscr{A}_B \wedge Acc)$, and by (K*1), $\psi^+ *_a (\varphi^+ \wedge \mathscr{A}_B \wedge Acc)$ is closed under classical logical consequence and hence contains all such α^+. By monotonicity of classical logic, it follows that $\psi^+ *_a (\varphi^+ \wedge \mathscr{A}_B \wedge Acc) \vdash \gamma^+$. By a similar argument, we can show that $\neg\psi^- *_a (\neg\varphi^- \wedge \mathscr{A}_B \wedge Acc) \vdash \neg\gamma^-$ and hence $\gamma \in \psi *_b \varphi$.

(K*2) $\varphi \in \psi *_b \varphi$.

Remember that by (K*2), $\varphi^+ \wedge \mathscr{A}_B \wedge Acc \in \psi^+ *_a (\varphi^+ \wedge \mathscr{A}_B \wedge Acc)$ and $\neg\varphi^- \wedge \mathscr{A}_B \wedge Acc \in \neg\psi^- *_a (\neg\varphi^- \wedge \mathscr{A}_B \wedge Acc)$. By (K*1), $\psi^+ *_a (\varphi^+ \wedge \mathscr{A}_B \wedge Acc) \vdash \varphi^+ \wedge \mathscr{A}_B \wedge Acc$ and hence $\psi^+ *_a (\varphi^+ \wedge \mathscr{A}_B \wedge Acc) \vdash \varphi^+$. Similarly, $\neg\psi^- *_a (\neg\varphi^- \wedge \mathscr{A}_B \wedge Acc) \vdash \neg\varphi^-$. By Definition 7.21, $\varphi \in \psi *_b \varphi$.

We argued in Section 2, page 270, that for non-classical logics, the two postulates (K*3) and (K*4) are equivalent to the following Postulate ($K^*_{3,4}$).

($K^*_{3,4}$) If A is consistent with K, then $K *_a A = \mathrm{Cn}_b(K \cup \{A\})$.

In this case, being 'consistent' means *acceptable*. In other words, that K and A together are acceptable,[7] in which case, the revision is equivalent to an expansion.

(K*5) $\psi *_b \varphi = \psi_\perp$ only if φ is contradictory.

Similarly, this should have the following interpretation '$\psi *_b \varphi$ is not acceptable, only if φ is (itself) not acceptable'. This can be shown if we remember that we artificially forced the translation of theories of $\mathbf{L}_B$ that are not acceptable to be classically inconsistent. This means, for instance, that if φ is itself acceptable, then $\psi^+ *_a (\varphi^+ \wedge \mathscr{A}_B \wedge Acc)$ is classically consistent.

Suppose $\psi *_b \varphi$ is not acceptable, we must show that φ is not acceptable. From the assumption, it follows that:

$$\psi *_b \varphi \rightarrow \gamma \text{ and } \psi *_b \varphi \rightarrow \neg\gamma, \tag{7}$$

for some γ in $\mathscr{L}_B$. By the correspondence theorem and Equation (7), $\{\mathscr{A}_B\} \cup \{\alpha^+ \mid \psi^+ *_a (\varphi^+ \wedge \mathscr{A}_B \wedge Acc) \vdash \alpha^+\} \vdash \gamma^+$.

Notice that $\{\mathscr{A}_B\} \cup \{\alpha^+ \mid \psi^+ *_a (\varphi^+ \wedge \mathscr{A}_B \wedge Acc) \vdash \alpha^+\} \subseteq \psi^+ *_a (\varphi^+ \wedge \mathscr{A}_B \wedge Acc)$ since by (K*2), $\mathscr{A}_B \in \psi^+ *_a (\varphi^+ \wedge \mathscr{A}_B \wedge Acc)$, and by (K*1), $\psi^+ *_a (\varphi^+ \wedge \mathscr{A}_B \wedge Acc)$ is closed under classical logical consequence and hence contains all such α^+. Also notice that by the correspondence theorem and Equation (7), $\{\mathscr{A}_B\} \cup \{\alpha^+ \mid \psi^+ *_a (\varphi^+ \wedge \mathscr{A}_B \wedge Acc) \vdash \alpha^+\} \vdash \overline{\gamma}^+$. For the same reason, $\{\mathscr{A}_B\} \cup \{\alpha^+ \mid \psi^+ *_a (\varphi^+ \wedge \mathscr{A}_B \wedge Acc) \vdash \alpha^+\} \subseteq \psi^+ *_a (\varphi^+ \wedge \mathscr{A}_B \wedge Acc)$ and therefore $\psi^+ *_a (\varphi^+ \wedge \mathscr{A}_B \wedge Acc) \vdash \gamma^+ \wedge \overline{\gamma}^+ = holds(\gamma, \mathsf{tt}) \wedge holds(\neg\gamma, tt)$.

Now $\mathscr{A}_B \wedge holds(\neg\gamma, tt) \vdash holds(\gamma, \mathsf{ff})$. By (K*1), $\psi^+ *_a (\varphi^+ \wedge \mathscr{A}_B \wedge Acc) \vdash holds(\gamma, \mathsf{tt}) \wedge holds(\gamma, \mathsf{ff})$. $Acc \in \psi^+ *_a (\varphi^+ \wedge \mathscr{A}_B \wedge Acc) \therefore \psi^+ *_a (\varphi^+ \wedge \mathscr{A}_B \wedge Acc) \vdash \perp$. By (K*5), $(\varphi^+ \wedge \mathscr{A}_B \wedge Acc)$ must be contradictory. $\mathscr{A}_B$ and Acc do not introduce themselves any inconsistencies, so it must be the case that φ is not itself acceptable.

(K*6) If $\varphi \equiv_b \gamma$, then $\psi *_b \varphi \equiv_b \psi *_b \gamma$.

[7] And consequently are acceptable separately as well.

https://doi.org/10.1017/S1755020308080246 Published online by Cambridge University Press

In $\mathbf{L}_B$, $\varphi \equiv_b \gamma$ means $\varphi \rightarrow \gamma$ and $\gamma \rightarrow \varphi$. By the correspondence: $\mathscr{A}_B \vdash holds(\varphi, \mathsf{tt}) \leftrightarrow holds(\gamma, \mathsf{tt})$ and also $\mathscr{A}_B \vdash \neg holds(\varphi, \mathsf{ff}) \leftrightarrow \neg holds(\gamma, \mathsf{ff})$. We show the positive part only. The negative part is similar. It is easy to see that $\vdash (\mathscr{A}_B \wedge \varphi^+) \leftrightarrow (\mathscr{A}_B \wedge \gamma^+)$. By (K*6), $\psi^+ *_a (\varphi^+ \wedge \mathscr{A}_B \wedge Acc) \equiv \psi^+ *_a (\gamma^+ \wedge \mathscr{A}_B \wedge Acc)$. The same happens with the negative part and therefore by the definition of revision in $\mathbf{L}_B$ $\psi *_b \varphi = \psi *_b \gamma$, and hence $\psi *_b \varphi \equiv_b \psi *_b \gamma$.

We argued in Section 2, page 5, that the two postulates (K*7) and (K*8) are equivalent to ($\mathrm{K}^*_{7,8}$) below.

($\mathrm{K}^*_{7,8}$) $\mathrm{Cn}_b(\psi *_b \varphi \cup \{\gamma\}) = \psi *_b (\varphi \wedge \gamma)$, when γ is consistent with $\psi *_b \varphi$.

In this case, being 'consistent' means *acceptable*. Acceptability plays therefore no role in the revision process above, and hence, $\psi *_b \varphi \cup \{\gamma\}$ is *acceptable*. In first-order logic terms, this means that the acceptability axiom is never used in the derivation of translated formulae of $\mathsf{wff}_{\mathscr{L}_B}$. We sketch the proof as follows. Let us look first at the formulae in the set $\psi *_b \varphi \cup \{\gamma\}$. These are the formulae in the set $\{\alpha \mid \psi^+ *_a (\varphi^+ \wedge \mathscr{A}_B \wedge Acc) \vdash \alpha^+$ and $\neg\psi^- *_a (\neg\varphi^- \wedge \mathscr{A}_B \wedge Acc) \vdash \neg\alpha^-\} \cup \{\gamma\}$. On the other hand, the formulae in the set $\psi *_b (\varphi \wedge \gamma)$ are the same formulae in the set $\{\beta \mid \psi^+ *_a ((\varphi \wedge \gamma)^+ \wedge \mathscr{A}_B \wedge Acc) \vdash \beta^+$ and $\neg\psi^- *_a (\neg(\varphi \wedge \psi)^- \wedge \mathscr{A}_B \wedge Acc) \vdash \neg\beta^-\}$. Now by the assumption and ($\mathrm{K}^*_{7,8}$) for $*_a$, we have that $\mathrm{Cn}(\psi^+ *_a (\varphi^+ \wedge \mathscr{A}_B \wedge Acc) \cup \{\gamma^+\}) = \psi^+ *_a ((\varphi \wedge \gamma)^+ \wedge \mathscr{A}_B \wedge Acc)$. Notice that $\mathscr{A}_B \in \psi^+ *_a (\varphi^+ \wedge \mathscr{A}_B \wedge Acc) \cup \{\gamma^+\}$. It is easy to show that $\psi^+ *_a (\varphi^+ \wedge \mathscr{A}_B \wedge Acc) \cup \{\gamma^+\} \vdash \beta^+$ and $\neg\psi^- *_a (\neg\varphi^- \wedge \mathscr{A}_B \wedge Acc) \cup \{\neg\gamma^-\} \vdash \neg\beta^-$ if and only if $\beta \in \mathrm{Cn}_b(\{\alpha \mid \psi^+ *_a (\varphi^+ \wedge \mathscr{A}_B \wedge Acc) \vdash \alpha^+$ and $\neg\psi^- *_a (\neg\varphi^- \wedge \mathscr{A}_B \wedge Acc) \vdash \neg\alpha^-\} \cup \{\gamma\})$. Similarly, $\psi^+ *_a ((\varphi \wedge \gamma)^+ \wedge \mathscr{A}_B \wedge Acc) \vdash \beta^+$ and $\neg\psi^- *_a (\neg(\varphi \wedge \gamma)^- \wedge \mathscr{A}_B \wedge Acc) \vdash \neg\beta^-$ if and only if $\beta \in \mathrm{Cn}_b(\{\alpha \mid \psi^+ *_a ((\varphi \wedge \gamma)^+ \wedge \mathscr{A}_B \wedge Acc) \vdash \alpha^+$ and $\neg\psi^- *_a (\neg(\varphi \wedge \gamma)^- \wedge \mathscr{A}_B \wedge Acc) \vdash \neg\alpha^-\})$.

8. Conclusions. We have presented a way of exporting an AGM revision process in classical logic to non-classical logics. We have illustrated our methodology by considering three types of logics. We started with the case of propositional modal logic whose notion of consistency is similar to that of classical logic; then we considered Łukasiewicz's many-valued logic, in which the notion of consistency is somewhat different but still exists; and finally, we investigated Belnap's four-valued logic that does not have a notion of consistency.

We have also outlined a general methodology for algebraic logics and shown that our approach can be used for any non-classical logic whose semantics is first-order logic axiomatisable.

We have shown that if the object logic has a notion of consistency that can be mapped into consistency in classical logic, then our approach can be used in a more or less straightforward way. This is because AGM revision is deeply connected with the notion of consistency in classical logic. Translations into classical logic of theories of object logics where the semantics is significantly different (e.g., paraconsistent logics) will always be consistent and hence revisions will amount to simple expansions.[8] However, the fact that

[8] This is in agreement with Tanaka (1997, p. 6) and Priest (2001, p. 219) who noted that some of the sphere semantics based on some paraconsistent logics show that revisions and expansions are equivalent.

https://doi.org/10.1017/S1755020308080246 Published online by Cambridge University Press

these logics do not trivialise in the face of 'inconsistency' does not mean that every belief set in these logics is 'rational', so a different notion replacing the task of consistency in classical logic needs to be employed. We have called this notion acceptability, and in our setting, the only requirement is that it can be expressed in classical logic. We do not claim to provide an analysis of what this notion should be in various non-classical scenarios. Our aim in this article was to provide a quick way to export a revision machinery in classical logic to these logics. For the case of Belnap's four-valued logic, we deemed acceptable those theories that do not prove a proposition and its negation. When applied in this context, our revision approach *filters out* all contradictions leaving only non-controversial information. This is much in the same spirit as the idea of 'locally restoring consistency, proposed by Restall & Slaney (1995), except that in our case, the 'vocabulary' whose beliefs are left untouched is restricted to the atoms for which contradictory beliefs are not held. As Restall and Slaney suggested, it is also possible to gradually restore consistency instead by defining successive notions of acceptability, which concentrate on particular sets of atoms at each time (also in the spirit of Fuhrmann, 1991).

There are considerable benefits for adopting our approach. Various translation mechanisms have been developed for many non-classical logics, which allow to mimic the non-classical logical reasoning within the context of classical logic. Existing results in the literature have already shown the benefits that such translations mechanisms can provide, for example, to give an object logic semantics (Broda *et al.*, 2000), to compare it with other logics, to get decidability/undecidability results, to make use of automated deduction of classical logic, and so forth. Our approach can equally benefit from building on a translation method and classical AGM belief revision, as it can allow to make use of existing automated systems for classical belief revision but in the context of non-classical logics. Moreover, revision theory for classical logic is very well developed. There are various notions and fine-tuning involved, and using translation mechanisms into classical logic will not only open a wealth of distinctions for the source logic but also enrich classical logic revision itself, with new ideas and problems arising from non-classical logics.

Furthermore, from the point of view of classical logic, our approach is a first example of a more general revision process, called *relative revision*. This concept can be defined as follows. Given a theory T of classical logic (e.g., a theory of linearly ordered Abelian groups), we can talk about a set of sentences Δ being acceptable (i.e., $\mathrm{Cn}(T \cup \Delta)$ being acceptable). We can talk about revision relative to T of Δ by ψ (denoted by $\Delta *_T \psi$) yielding a new acceptable theory $\Delta' = \Delta *_T \psi$. The case where T arises in connection with a translation from another logic is only one instance of this general relative revision. In Gabbay *et al.* (2006), we have investigated this problem in the context of the voting problem. A complete and detailed framework for relative revision is subject of our future work.

9. Discussion. This article is part of the general theme *logic by translation*. Given a logic L_1 and another logic L_2, then we can learn properties of the logic L_1 if we can translate L_1 faithfully and soundly into L_2 and if we have better knowledge of the properties of L_2. The most common logic L_2 to use is classical logic.

Take the case of interpolation, a property of great interest and many applications for any logic. Assume that we are dealing with a logic L_1 and we want to know whether it has interpolation. So let $\varphi(p, q)$ and $\psi(q, r)$ be two formulae of L_1 with the indicated atomic variables. We are seeking an interpolant $\theta(q)$ with the common atom q such that $\varphi(p, q) \vdash_{L_1} \theta(q) \vdash_{L_1} \psi(q, r)$. Such an interpolant may not exist for L_1.

https://doi.org/10.1017/S1755020308080246 Published online by Cambridge University Press

Let us now do *interpolation by translation*. Translate φ and ψ into L_2. Let φ^* and ψ^* be the translations. Usually L_2 is classical logic and the translation $*$ is sound and faithful, so we also have that $\varphi^* \vdash_{L_2} \psi^*$. If we have interpolation in L_2, then we have in L_2 an interpolant I such that $\varphi^* \vdash_{L_2} I \vdash_{L_2} \psi^*$.

We now have to look for a θ of L_1 such that $\theta^* \iff I$. If we find it, we get $\varphi \vdash_{L_1} \theta \vdash_{L_1} \psi$. The problem is how to find such a θ. In other words, we may not be able to translate back from L_2 to L_1. When there is no interpolation for L_2, the translation suggests how to turn L_2 into a stronger L_2^+, so that interpolants for L_2 exist in L_2^+ (Gabbay & Maksimova, 2005, Ch. 14).

Another common use of *logic by translation* is finding semantics for L_1. If we have semantics for L_2, we immediately get semantics for L_1. The translation $*$ may be obtained in a syntactical manner, using a proof theoretical interpretation. There are translations of classical logic into intuitionistic logic and translations of intuitionistic logic into modal logic, so any new semantics for modal logic would give a new semantics for intuitionistic and classical logics. Anything we do with the possible worlds, semantics of modal logic can be exported to intuitionistic and classical logics. For example, there are many proposals for endowing modal logic with probability. This may give new ideas for probability on other logics.

Another example of successful translations into modal logic include Boutilier's translation of the conditional. There are many more.

To summarise, *revision by translation* is part of a wider methodology of *logic by translation* and can give and suggest quick general results for any new logic L_1 that can be translated into well-known logics. Of course, one may wish to investigate in more detail the logic L_1 directly, maybe even using the translations of L_1 into $L_2^1, L_2^2, \ldots$, as sources of clues as to what the options may be.

10. Acknowledgement. We thank the referees for their comments.

BIBLIOGRAPHY

Alchourrón, C. E., Gärdenfors, P., & Makinson, D. (1985). On the logic of theory change: partial meet contraction and revision functions. *Journal of Symbolic Logic*, **50**(2), 510–530.

Anderson, A. R., Belnap, N. D., & Dunn, J. M. (1992). *A Useful Four-Valued Logic: How a Computer Should Think* [Section 81 of Anderson, A. R., and Belnap, N. D., Jr., and Dunn, J. M., *et al. Entailment*, vol. II, pp. 506–541]. Princeton, NJ: Princeton University Press.

Belnap, N. (1977a). How a computer should think. In Ryle, G., editor, *Contemporary Aspects of Philosophy*. Stocksfield: Oriel Press, pp. 30–56.

Belnap, N. (1977b). A useful four-valued logic. In Dunn, J. M. and Epstein, G., editors, *Modern Uses of Multiple-Valued Logic*. Stocksfield: Reidel, pp. 8–37.

Broda, K., Russo, A., & Gabbay, D. (2000). Discovering world with fuzzy logic: perspectives and approaches to the formalisation of human-consistent logical systems. In Novak, V., and Perf, I., editors, *A Unified Compilation Style Labelled Deductive Systems for Modal, Substructural and Fuzzy Logic*. Berlin, Germany: Springer-Verlag, pp. 495–547.

Chellas, B. F. (1980). *Modal Logic: an Introduction*. Cambridge, UK: Cambridge University Press.

https://doi.org/10.1017/S1755020308080246 Published online by Cambridge University Press

Darwiche, A., & Pearl, J. (1996). On the logic of iterated belief revision. Technical Report R-202. Cognitive Science Laboratory, Computer Science Department, University of California, Los Angeles, CA.

Darwiche, A., & Pearl, J. (1997). On the logic of iterated belief revision. *Artificial Intelligence*, **89**, 1–29.

Fuhrmann, A. (1991). Theory contraction through base contraction. *Journal of Philosophical Logic*, **20**(2), 175–203. DOI 10.1007/BF00284974.

Gabbay, D. M., & Maksimova, L. (2005). *Interpolation and Definability. Modal and Intuitionistic Logics*, vol. 1. Oxford, UK: Oxford Science Publications. ISBN 0-19-851174-4.

Gabbay, D. M., Pigozzi, G., & Rodrigues, O. (2006). Belief revision, belief merging and voting. In Bonanno, G., van der Hoek, W., and Wooldridge, M., editors, *Proceedings of the Seventh Conference on Logic and the Foundations of Games and Decision Theory (LOFT06)*. Liverpool, UK: University of Liverpool, pp. 71–78.

Gabbay, D., Rodrigues, O., & Russo, A. (2000). Information, uncertainty, fusion. In Yager, R. R., Zadeh, L. A., and Bouchon-Meunier, B., editors, *Revision by Translation*. Norwell, MA: Kluwer Academic Publishers, pp. 3–31. ISBN: 0-7923-8590-X.

Gärdenfors, P. (1988). *Knowledge in Flux: Modeling the Dynamics of Epistemic states. A Bradford Book*. Cambridge, MA: The MIT Press.

Katsuno, H., & Mendelzon, A. O. (1991). Propositional knowledge base revision and minimal change. *Artificial Intelligence*, **52**, 263–294.

Martins, J. P., & Shapiro, S. C. (1988). A model for belief revision. *Artificial Intelligence*, **35**(1), 25–79.

Ohlbach, H. J. (1991). Semantics-based translations methods for modal logics. *Journal of Logic and Computation*, **1**(5), 691–746.

Priest, G. (2001). Paraconsistent belief revision. *Theoria*, **67**, 214–228.

Restall, G., & Slaney, J. (1995). Realistic belief revision. In De Glas, M., and Pawlak, Z., editors, *Wocfai 95: Proceedings of the Second World Conference on the Fundamentals of Artificial Intelligence*, Angkor, July 1995. Paris, France: Angkor, pp. 67–378. ISBN 2-87892-009-0.

Shapiro, S. C. (1992). *Relevance logic in computer science* [Section 83 of Anderson, A. R., and Belnap, N. D., Jr., and Dunn, J. M., *et al. Entailment*, vol. II, pp. 553–563]. Princeton, NJ: Princeton University Press.

Tanaka, K. (1997). *What Does Paraconsistency Do?* Prague, Czech Republic: Filosofia. *The Case of Belief Revision. The Logica Yearbook*.

DEPARTMENT OF COMPUTER SCIENCE,
KING'S COLLEGE LONDON
LONDON WC2R 2LS, UK
E-mail: dov.gabbay@kcl.ac.uk

DEPARTMENT OF COMPUTER SCIENCE,
KING'S COLLEGE LONDON
LONDON WC2R 2LS, UK
E-mail: odinaldo.rodrigues@kcl.ac.uk

DEPARTMENT OF COMPUTING,
IMPERIAL COLLEGE,
LONDON SW7 2BZ, UK
E-mail: ar3@doc.ic.ac.uk

https://doi.org/10.1017/S1755020308080246 Published online by Cambridge University Press