DEPTH RELEVANCE AND HYPERFORMALISM

SHAY ALLEN LOGAN

Abstract. Formal symptoms of relevance usually concern the propositional variables shared between the antecedent and the consequent of provable conditionals. Among the most famous results about such symptoms are Belnap's early results showing that for sublogics of the strong relevant logic R, provable conditionals share a signed variable between antecedent and consequent.

For logics weaker than R stronger variable sharing results are available. In 1984, Ross Brady gave one well-known example of such a result. As a corollary to the main result of the paper, we give a very simple proof of a related but strictly stronger result.

NOTE: This is an updated version of a paper published in the *Journal of Philosophical Logic*. The published version contains an error pointed for which there is published errata. This version does not contain the error.

1. Introduction

Intuitively, the logic L is a *relevant* logic just when $\vdash_L A \to B$ only if the content of A is somehow related to the content of B. Formal symptoms of relevance usually concern the propositional variables shared between the antecedent and the consequent of provable conditionals. Among the most famous results about such symptoms are the following theorems proved in 1960 by Belnap.[1]

Theorem 1 (The Weak Belnap Theorem)**.** *If L is a sublogic of the logic R and $\vdash_L A \to B$, then some variable p occurs in both A and B.*

Theorem 2 (The Strong Belnap Theorem)**.** *If L is a sublogic of the logic R and $\vdash_L A \to B$, then some variable p occurs with the same sign in both A and B.*

Note that the notion of 'sign' invoked in Theorem 2 will be explained below. For the moment, we note only that for logics weaker than R, we can strengthen the conditions we put on the variable being shared. One step in this direction was taken by Brady in 1984, when he proved the following result:

Theorem 3 (The Weak Brady Theorem)**.** *If L is a sublogic of the logic DR and $\vdash_L A \to B$, then some variable p occurs at the same depth in both A and B.*

In [5], using essentially the same techniques Brady used, I proved the following strengthening of his result:

Theorem 4 (The Strong Brady Theorem)**.** *If L is a sublogic of the logic DR and $\vdash_L A \to B$, then some variable p occurs at the same depth and with the same sign in both A and B.*

[1] See [2] for Belnap's proof of what I here am calling the Strong Belnap Theorem. Since the Weak Belnap Theorem clearly follows from the Strong Belnap Theorem, we should take this to be the first instance of a proof of either result. Also worth noting is that the proof in [2] only technically proves the weak Belnap theorem for subsystems of the weaker logic E, rather than R. However, the proof in [2] is taken up nearly word-for-word as the proof of Theorem 2 found in Section 22.1.3 of [1].

What we will show in this paper is that, for a restricted logic that in [**?**] I called DR$^-$, the Belnap Theorems and the Brady Theorems are more tightly connected than meets the eye. In particular one consequence of the result I prove is that logics generated by any subset of DR's axioms using any subset of DR$^-$'s rules have a feature that I below call *hyperformality*, and it is a consequence of hyperformality that, in such logics, violations of the Weak Brady Theorem can be turned into violations of the Weak Belnap Theorem and violations of the Strong Brady Theorem can be turned into violations of the Strong Belnap Theorem. A nice consequence of this discussion is that we get nice, 'user-friendly' proofs of Brady's results, restricted to these logics.

2. Some Setup

We work in a standard propositional language with the connectives $\wedge$, $\vee$, $\rightarrow$, and $\neg$. Our interest will center on the strong relevant logic R made famous by [1], the logic called DR in [3], and the fragment of DR known as DR$^-$ examined in [**?**]. These logics contain the logic DW, axiomatized below, as a common fragment:

A1. $A \rightarrow A$

A2. $(A \wedge B) \rightarrow A/B$

A3. $A/B \rightarrow (A \vee B)$

A4. $((A \rightarrow B) \wedge (A \rightarrow C)) \rightarrow (A \rightarrow (B \wedge C))$

A5. $((A \rightarrow C) \wedge (B \rightarrow C)) \rightarrow ((A \vee B) \rightarrow C)$

A6. $(A \wedge (B \vee C)) \rightarrow ((A \wedge B) \vee (A \wedge C))$

A7. $\neg\neg A \rightarrow A$

A8. $(A \rightarrow \neg B) \rightarrow (B \rightarrow \neg A)$

R1. $A, B \Rightarrow A \wedge B$

R2. $A, A \rightarrow B \Rightarrow B$

R3. $A \rightarrow \neg B \Rightarrow B \rightarrow \neg A$

R4. $A \rightarrow B, C \rightarrow D \Rightarrow (B \rightarrow C) \rightarrow (A \rightarrow D)$

DR extends DW with the following two axioms and three rules:

A9. $((A \rightarrow B) \wedge (B \rightarrow C)) \rightarrow (A \rightarrow C)$

A10. $A \vee \neg A$

R4. $C \vee A, C \vee (A \rightarrow B) \Rightarrow C \vee B$

R5. $C \vee A \Rightarrow C \vee \neg(A \rightarrow \neg A)$

R6. $E \vee (A \rightarrow B), E \vee (C \rightarrow D) \Rightarrow E \vee ((B \rightarrow C) \rightarrow (A \rightarrow D))$

DR$^-$, in contrast, adds only the following rule to DW:

R7. $A \Rightarrow \neg(A \rightarrow \neg A)$

R instead adds these three axioms to DW:

A11. $(A \rightarrow B) \rightarrow ((B \rightarrow C) \rightarrow (A \rightarrow C))$

A12. $A \rightarrow ((A \rightarrow B) \rightarrow B)$

A13. $(A \rightarrow (A \rightarrow B)) \rightarrow (A \rightarrow B)$

It is a straightforward exercise to verify that all the axioms and rules of DR are derivable in R, from which it follows that DR (and thus DR^-) is a sublogic of R.

3. The Belnap Theorems

To begin, we recursively define the *sign* of an occurrence of a subformula as follows:

- Every formula occurs positively as a subformula of itself.
- If an occurrence of $B \wedge C$ is positive, then the corresponding occurrences of B and C are positive as well.
- If an occurrence of $B \wedge C$ is negative, then the corresponding occurrences of B and C are negative as well.
- If an occurrence of $B \vee C$ is positive, then the corresponding occurrences of B and C are positive as well.
- If an occurrence of $B \vee C$ is negative, then the corresponding occurrences of B and C are negative as well.
- If an occurrence of $B \rightarrow C$ is positive, then the corresponding occurrence of B is negative and the corresponding occurrence of C is positive.
- If an occurrence of $B \rightarrow C$ is negative, then the corresponding occurrence of B is positive and the corresponding occurrence of C is negative.
- If an occurrence of $\neg B$ is positive, then the corresponding occurrence of B is negative.
- If an occurrence of $\neg B$ is negative, then the corresponding occurrence of B is positive.

If you prefer, you can picture these rules as acting on parse trees using the following rules:

ISPR: $\underset{+}{A}$

$\wedge_+$**SPR:** $A \underset{+}{\wedge} B \Rightarrow \underset{+}{A} \wedge \underset{+}{B}$

$\wedge_-$**SPR:** $A \underset{-}{\wedge} B \Rightarrow \underset{-}{A} \wedge \underset{-}{B}$

$\vee_+$**SPR:** $A \underset{+}{\vee} B \Rightarrow \underset{+}{A} \vee \underset{+}{B}$

$\vee_-$**SPR:** $A \underset{-}{\vee} B \Rightarrow \underset{-}{A} \vee \underset{-}{B}$

$\rightarrow_+$**SPR:** $A \underset{+}{\rightarrow} B \Rightarrow \underset{-}{A} \rightarrow \underset{+}{B}$

$\rightarrow_-$**SPR:** $A \underset{-}{\rightarrow} B \Rightarrow \underset{+}{A} \rightarrow \underset{-}{B}$

$\neg_+$**SPR:** $\underset{+}{\neg} A \Rightarrow \neg \underset{-}{A}$

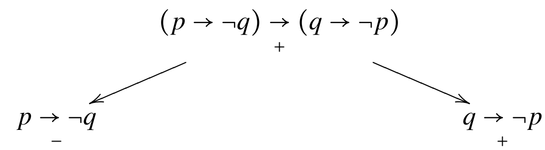
$\neg_-$**SPR:** $\underset{-}{\neg} A \Rightarrow \neg \underset{+}{A}$

We read these as follows: ISPR—the Initial Sign Parse Rule—says to begin by placing a plus under the formula being examined. $\wedge_+$SPR says that a plus placed under conjunction should lead, in the next 'level' of the parse tree, to plusses under each conjunct. The remaining rules are interpreted similarly. Both to show the rules in action and to address a common misunderstanding, we'll pause to look at an example.

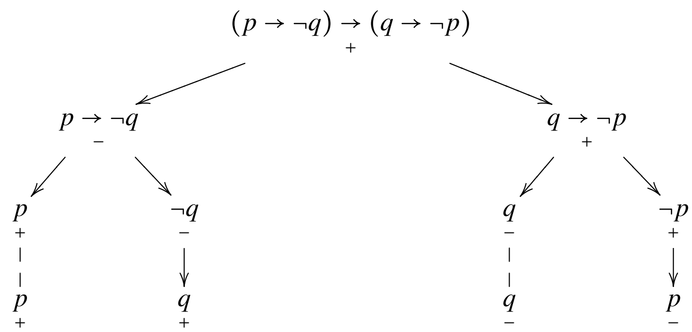
Example 1: Consider the formula $(p \to \neg q) \to (q \to \neg p)$. If we follow the parse-tree rules above, ISPR tells us to begin with the following:

$$(p \to \neg q) \underset{+}{\to} (q \to \neg p)$$

We then apply the appropriate parse rule—in this case, $\to_+$SPR. The next step thus looks like this:



Continuing this process, the end result is the following labeled tree:



Where here the dashed lines simply repeat prior information so the end result is easier to read. End Example

Recall that the Strong Belnap Theorem said that if L is a sublogic of the logic R and $\vdash_L A \to B$, then some variable p occurs with the same sign in both A and B. It is easy to misunderstand the content of this theorem in a way that makes it look like Example 1 gives a counterexample. After all, the formula we began with clearly is a theorem—indeed, an axiom!—of R. But as the parse tree makes clear, *every* atom that occurs in the antecedent is positive in its occurrence there and *no* atom that occurs in the consequent is positive in its occurrence there. And that sure *seems* like a clear violation of the Strong Belnap Theorem.

To see what's gone wrong notice that the Strong Belnap Theorem concerns the signs of atomic subformulas of the antecedent *as subformulas of the antecedent* and of atomic subformulas of the consequent *as subformulas of the consequent*. But the sign of an occurrence of a given formula A as a subformula of some formula B will, in general, be different than the sign of the same occurrence of A when considered as a subformula of some formula B' that B is itself a subformula of. In the case at hand, what matters is that the sign of an occurrence of a variable p as a subformula of A will in general be different than the sign of that same occurrence of p as a subformula of $A \to B$.

Differences in this neighborhood will cause a bit of unavoidable confusion later, so let's pause to make sure the point is perfectly clear. The basic point is that Theorems 1, 2, 3, and 4 all have roughly the following form:

> For appropriate logics L, if $\vdash_L A \to B$, then the labels attached to the parse tree for A and the labels attached to the parse tree for B have thus and such in common.

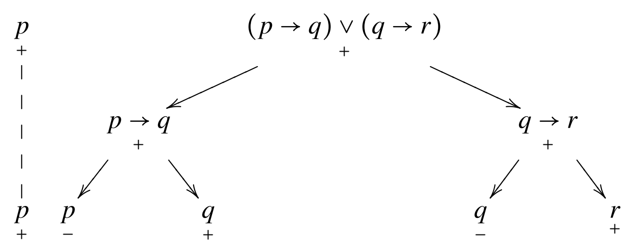
What they *don't* say is anything that has this form:

> For appropriate logics L, if $\vdash_L A \to B$, then in the parse tree for $A \to B$, the labels attached to the subtree under A and the labels attached to the subtree under B have thus and such in common.

All told, what this makes clear is that the right parse trees to look at are these ones:

And now the apparent counterexample vanishes. Indeed, now *every* variable that occurs in the antecedent occurs with the same sign in the consequent.

The Belnap Theorems are really quite useful. To see this, let's have a look at another example:

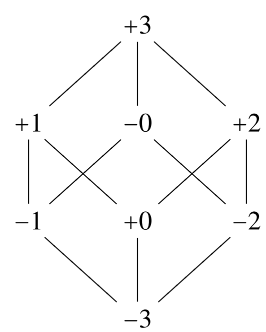
> Example 2: Since $(p \to q) \vee (q \to r)$ is a theorem of classical logic, the following is as well:
>
> $$p \to ((p \to q) \vee (q \to r))$$
>
> But anyone with even a hint of relevant scruples ought to hope that it isn't a theorem of R. And, indeed, it isn't. But in any of the usual model theoretic semantic theories for R, providing a countermodel would be a nontrivial—not to mention annoying—exercise. Luckily, the parse-tree method of checking for strong variable sharing is quick and painless. For convenience, here are the necessary parse trees:
>
> 
>
> And there we have it: no signed variable is shared from antecedent to consequent. Thus $p \to ((p \to q) \vee (q \to r))$ is not a theorem of R or any of its sublogics. End Example

In spite of their usefulness, however, the Belnap Theorems really do only manage to give necessary and not sufficient conditions for R-theoremhood. For example, $p \to (q \to p)$ is canonically a non-theorem of R. But it *does* have a positive occurrence of p in both antecedent and consequent. Thus, while *failing to share* a (signed) variable rules *out* a conditional as a theorem of R, *sharing* a (signed) variable doesn't rule *in* a conditional as a theorem of R.

We'll now turn to thinking about Belnap's proof of the Strong Belnap Theorem. It has roughly the following structure:

- First, he defines a class of interpretations of the language that take values in certain lattice known as M_0.
- With the interpretations on hand, he then proves two key lemmas:
  (1) All Theorems of R are assigned a designated value on each interpretation.
  (2) If A and B don't share a signed variable, then there is an interpretation that assigns $A \to B$ a non-designated value.
- Together, the lemmas show that if A and B don't share a signed variable, $A \to B$ isn't a theorem of R.
- Contraposing finishes the job.

This is an argument structure logicians should, generally speaking, feel comfortable with. What's more is that the two main characters here—the lattice M_0 and the class of interpretations of our language into M_0—are themselves pretty friendly. More to the point, M_0 has the following Hasse diagram.



This lattice is very familiar from the algebraic study of relevant logics. In fact, when Anderson and Belnap first introduce M_0 in Entailment Volume 1, they say the following:

> Stone 1936 showed that a maximal filter of a Boolean algebra B determines a homomorphism of B into the particularly simple two-element Boolean algebra. We are about to prove an analogous theorem for intensional lattices, but first we must characterize the particular simple intensional lattice, which we call M_0, that is used in our theorem.[1, §18.4]

The class of interpretations Belnap actually gives are also quite nice. To begin, say that an *assignment* is a function v that maps propositional variables into M_0. Belnap then extends each assignment to a function v^+ mapping arbitrary formulas into M_0 by the following—pretty vanilla looking—recursive clauses:

- $v^+(\neg A) = -1 \cdot v^+(A)$;
- $v^+(A \wedge B) = \inf(v^+(A), v^+(B))$;
- $v^+(A \vee B) = \sup(v^+(A), v^+(B))$;
- $v^+(A \to B) = m_\to(v^+(A), v^+(B))$ where $m_\to$ has the following matrix:

y									
$+3$	$+3$	$+3$	$+3$	$+3$	$+3$	$+3$	$+3$	$+3$	
$+2$	$+3$	$+2$	-3	-3	$+2$	-3	$+2$	-3	
$+1$	$+3$	-3	$+1$	-3	$+1$	$+1$	-3	-3	
$+0$	$+3$	-3	-3	-3	$+0$	-3	-3	-3	
-0	$+3$	$+2$	$+1$	$+0$	-0	-1	-2	-3	
-1	$+3$	-3	$+1$	-3	-1	-1	-3	-3	
-2	$+3$	$+2$	-3	-3	-2	-3	-2	-3	
-3	$+3$	-3	-3	-3	-3	-3	-3	-3	
$m_{\rightarrow}(x,y)$	-3	-2	-1	-0	$+0$	$+1$	$+2$	$+3$	x

$m_{\rightarrow}$ is not the weirdest function to interpret the arrow, but I do think it stands in need of some justification.

To provide such justification, it helps to think of the elements of M_0 as having the form σn where $\sigma \in \{+,-\}$ and $n \in \{0,1,2,3\}$, and to interpret these values in the following way:

- 'formula A having value $+3$' is interpreted as 'A is totally true'; A having other positive values is interpreted as other ways for A to be (less than totally) true.
- 'formula A having value -3' is interpreted as 'A is totally false'; A having other negative values is interpreted as other ways for A to be (less than totally) true.

With all that on board, here are two observations. First, $m_{\rightarrow}$ is *almost* the function defined as follows:

$$m^*_{\rightarrow}(\sigma_1 a, \sigma_2 b) = \begin{cases} +\max(a,b) & \text{if } \sigma_2 b \geq \sigma_1 a \\ -\max(a,b) & \text{if } \sigma_2 b < \sigma_1 a \\ -3 & \text{otherwise} \end{cases}$$

Second, $m^*_{\rightarrow}$ can intuitively be understood to be evaluating conditionals in the following way:

- $A \rightarrow B$ is as false as possible when A and B can't be compared. Otherwise,
- $A \rightarrow B$ is as close to being totally true/totally false as the closer of its two arguments; and
  - $A \rightarrow B$ is true when it should be—that is, when we don't decrease in truth value when we pass from A to B; and
  - $A \rightarrow B$ is false otherwise—that is, when we do decrease in truth value as we pass from A to B.

This is an intelligible and not immediately ridiculous way to evaluate the truth of conditionals. The resulting matrix has the following form, where I've underlined the nine entries that are different from the corresponding entries in $m_{\rightarrow}$:

y									
$+3$	$+3$	$+3$	$+3$	$+3$	$+3$	$+3$	$+3$	$+3$	
$+2$	$+3$	$+2$	-3	-3	$+2$	-3	$+2$	-3	
$+1$	$+3$	-3	$+1$	-3	$+1$	$+1$	-3	-3	
$+0$	$+3$	-3	-3	-3	$+0$	$\underline{-1}$	$\underline{-2}$	-3	
-0	$+3$	$+2$	$+1$	$+0$	$\underline{-3}$	$\underline{-3}$	$\underline{-3}$	-3	
-1	$+3$	-3	$+1$	$\underline{-1}$	$\underline{-3}$	-1	-3	-3	
-2	$+3$	$+2$	-3	$\underline{-2}$	$\underline{-3}$	-3	-2	-3	
-3	$+3$	-3	-3	-3	-3	-3	-3	-3	
$m^*_{\rightarrow}(x,y)$	-3	-2	-1	-0	$+0$	$+1$	$+2$	$+3$	x

So Belnap's interpretation of the conditional is *almost* the intuitive understanding just given. But the difference is needed. For example, suppose v takes the following values: $v(p) = +1$,

$v(q) = -1$, and $v(r) = +0$. If we compute $v^+(((p \to q) \wedge (p \to r)) \to (p \to (q \wedge r)))$ using $m^*_\to$ to interpret the arrow, the result (as we leave the reader to check) is -3 which is *not* a designated value. But if we evaluate it using $m^\to$, as Belnap suggested, we instead get $+3$, which is.

Altogether the point is this: Belnap's proof of the Strong Belnap theorem has a familiar form. The characters inhabiting the proof are also mostly familiar. The one exception—the function $m_\to$—isn't so bad, being a mere tweak away from something we can make sense of.

4. The Brady Theorems

We'll now turn to looking at the Brady Theorems. For convenience, we begin by restating them:

Weak Brady Theorem: If L is a sublogic of the logic DR and $\vdash_L A \to B$, then some variable p occurs at the same depth in both A and B.

Strong Brady Theorem: If L is a sublogic of the logic DR and $\vdash_L A \to B$, then some variable p occurs at the same depth and with the same sign in both A and B.

The notion of *depth* deployed here is defined as follows::

- A occurs at depth 0 in A.
- Given a depth n occurrence of $\neg A$ in C, the corresponding occurrence of A in C is a depth n occurrence as well.
- Given a depth n occurrence of $A \wedge B$ in C, the corresponding occurrences of A and of B in C are depth n occurrences as well.
- Given a depth n occurrence of $A \vee B$ in C, the corresponding occurrences of A and of B in C are depth n occurrences as well.
- Given a depth n occurrence of $A \to B$ in C, the corresponding occurrences of A and of B in C are depth $n + 1$ occurrences.

Again we can see these as parse tree rules. Stated using the same notation (and with essentially the same reading) as before, they look like this:

- $\underset{0}{A}$
- $\underset{n}{\neg} A \Rightarrow \neg \underset{n}{A}$
- $A \underset{n}{\wedge} B \Rightarrow \underset{n}{A} \wedge \underset{n}{B}$
- $A \underset{n}{\vee} B \Rightarrow \underset{n}{A} \vee \underset{n}{B}$
- $A \underset{n}{\to} B \Rightarrow \underset{n+1}{A} \to \underset{n+1}{B}$

Either way we look at it, the important part is that depth measures how '$\to$-nested' a subformula is.

At a casual glance, The Strong Brady Theorem might seem like it's nothing more than the conjunction of the Weak Brady Theorem and the restriction of the Strong Belnap Theorem to sublogics of DR. But this is incorrect: the Strong Brady Theorem is in fact *strictly stronger* than this conjunction. To see this, note that if we conjoin the Weak Brady Theorem and the restricted Strong Brady Theorem, the result is a theorem of the following form:

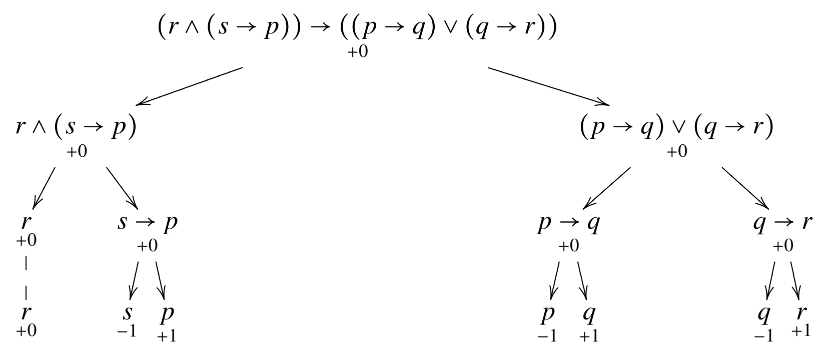
> Whenever L is a sublogic of DR and $\vdash_L A \to B$, some variable p occurs at the same depth in both A and B and some variable q occurs with the same sign in both A and B.

And while this is a *consequence* of the Strong Brady Theorem, the Strong Brady Theorem guarantees strictly more; in particular, it guarantees that the very same variable occurs with (simultaneously) the same depth and the same sign in both A and B. This strengthening turns out to be useful, as we will make clear by examining an example. As a warning, we will have reason to think about this very same example later in the paper.

Example 3: Consider the following formula:

$$(r \wedge (s \to p)) \to ((p \to q) \vee (q \to r))$$

Note that, for the same reasons as in Example 2, this is a theorem of classical logic. If we apply both parse tree procedures at once in a natural way, the end result is the following tree:



From this we can see that the formula *does* share a signed variable between antecedent and consequent because r has a positive occurrence in each. And it *does* share a variable at the same depth between antecedent and consequent because p has a depth 1 occurrence in each. But no variable has an occurrence that is simultaneously at the same depth and of the same sign in both antecedent and consequent.

Thus while neither the Strong Belnap Theorem nor the Weak Brady Theorem nor their conjunction will rule this theorem of classical logic out as a theorem of DR, the Strong Brady Theorem does. End Example

That's not to say that Strong Brady is somehow perfect. It's not—just as with all the other theorems we're discussing, it gives a *necessary but not sufficient* condition for membership in the class of theorems in question. For example, if we switch s and p in the antecedent of the formula in Example 3, the resulting formula remains a theorem of classical logic and *does* have a negative, depth 1 occurrence of p in both antecedent and consequent despite not being a theorem of R (or of any of its sublogics).

So the Brady Theorems—and especially the Strong Brady Theorem—while still not perfect, are nonetheless useful. It's thus unfortunate that the proofs we have of them are not as nice as the proofs we have of the Belnap Theorems. In particular, the proofs we have rely on evaluating formulas not in individual assignments of values to variables, but in slightly more than infinitely long sequences ($\omega + 1$ sequences, to be precise) of assignments of values to variables, and then relying on something vaguely supervaluation-looking to wrap them all up into a single value.

However, because it's not necessary for the paper to be self-contained, and because it would take up to much space, I won't go into any more detail here. The interested reader is instead referred to [3] and [5]. In the remainder of the paper, I instead provide an alternative, and much simpler proof of a result to which the Strong Brady Theorem (and thus the Weak

Brady Theorm as well) is a straightforward corollary. The proof I give has another virtue as well: it shows that the Brady Theorems are, in a sense we can make precise, a natural extension of the Belnap Theorems to the setting of sufficiently weak logics.

5. Depth Substitutions

Our initial inspiration comes from an observation Brady makes in the course of proving the Weak Brady Theorem:

> The depth relevance condition suggests that there are levels of implication corresponding to the depths of sentential variable occurrences in a formula and that implication differs from level to level.[3, p. 65]

Personally, I find it hard to see either Brady Theorem showing this. The Weak Brady Theorem says that when $A \to B$ is provable, A and B share a variable at some depth. According to the Strong Brady Theorem, we can take this variable to have the same sign in both occurrences. I'll grant that these say *something* about how depth and implication interact. But so far as I can tell, they say nothing about there being different depths or 'levels' of implication.

For my money, for a system to recognize different 'levels of implication', it would need to be the case that if some formula C occurs as a subformula of A twice, once at one depth and once at some other depth, then these two occurrences of C would have to be seen—or at least 'seen for the purposes of implication—as occurrences not of the same formula at different depths, but in some sense as occurrences of two entirely different formulas. This is clearly a *very* loose idea, not least because I'm not all sure what it means for a logic to "see" things as different or the same. But, taken to the extreme, this loose idea becomes usable.

What I have in mind is this: rather than talking about different occurrences of an *arbitrary* subformula C, let's focus on the simplest and most fundamental case: occurrences of *variables*. Here the loose notion outlined above becomes this: a system recognizes different 'levels of implication' when any time some variable p occurs as a subformula of A at two different depths, these two occurrences of p are seen as occurrences of different formulas. This, like I said, is something I think we can make sense of. We'll need a few definitions first, beginning somewhere quite familiar:

- A uniform substitution is a function mapping each variable to a formula.
- If f is a uniform substitution and A is a formula, $A[f]$ is the formula that we get by replacing each occurrence of p in A with an occurrence of $f(p)$.

Typically, we require *formal* logics to be closed under uniform substitutions in the sense that whenever we have that $\vdash_L A$, we also have that $\vdash_L A[f]$ for all uniform substitutions f. The basic thought to have is this: we require this of formal logics because we require them to 'see' variables as schematic formulae.[2]

Thus, intuitively, a logic that 'sees' an occurrence of the same variable at two different depths as occurrences of two different variables should, by analogy, allow substitutions that vary depending on depth. To make this concrete we need a couple more definitions:

- A *depth substitution* is a function mapping pairs $\langle p, n\rangle$ with p a variable and $n \in \mathbb{N}$ to formulas.
- If d is a depth substitution and A is a formula, $A[d]$ is the formula that we get by replacing each depth-n occurrence of p with an occurrence of $d(p,n)$.

[2]This plausibly follows from taking logics to be either 2-formal or 3-formal in the sense of [6]. Given the content of what I'm *about to say*, there's clearly something very interesting and worthwhile to investigate here. Equally clear, however, is that such an investigation is beyond the scope of this particular paper.

Following the convention above, say that a set of formulas S is *closed under depth substitutions* when for any depth substitution d, if $\phi \in S$ then $\phi[d] \in S$ as well. Note that each uniform substitution f is naturally seen as a degenerate depth substitution d_f where we take $d_f(p,n) = f(p)$ for all n. Thus, any set of formulas that is closed under depth substitutions is also closed under uniform substitutions. Since closure under uniform substitution is an analogue of *formality*, we will call any logic that is closed under the more inclusive class of *depth* substitutions a *hyperformal* logic.

A rather unexpected result is that an enormous array of weak relevant logics, including DR^- and the sublogics of DR^- that are typically considered in the literature, are hyperformal. To be a bit more concrete, we make the following observations.

Lemma 5. *Let S be a set of formulas and S^+ be the smallest set that contains S and also contains the conclusion of each instance of R1-R6 whose premises it also contains (in other words, let S^+ be the closure of S under R1, R2, R3, R4, and R7). Then if S is closed under depth substitutions, so is S^+.*

Proof. If $B^+ \in S^+$, then for some $n \geq 0$, B^+ is the result of applying n instances of the rules in question to some $B \in S$. We prove by induction on n that for any depth substitution d, $B^+[d] \in S^+$.

The result holds by assumption for $n = 0$. The induction step splits into six cases depending on the last rule applied. We consider only the R2 (modus ponens) case, the others being either similar or simpler. So, suppose the last instance of a rule we applied was the following:

$$A, A \to B^+ \Rightarrow B^+$$

Let d be a depth substitution. Our goal is to show $B^+[d] \in S^+$. Define the depth substitution $d-1$ as follows:

$$(d-1)(p,n) = \begin{cases} p & \text{if } n = 0 \\ d(p,n-1) & \text{otherwise} \end{cases}$$

Note that $(A \to B^+)[d-1] = A[d] \to B^+[d]$. By the inductive hypothesis, for any depth substitutions e and e', $A[e] \in S^+$ and $(A \to B^+)[e'] \in S^+$. Thus in particular $A[d]$ and $(A \to B^+)[d-1] = A[d] \to B^+[d]$ are in S^+. So $B^+[d] \in S^+$ as required. □

From this we have the following corollary:

Corollary 6. *Let L be a logic generated from the set of axioms* Ax *using the set of rules* Ru*. Then if* Ax *is closed under depth substitutions and* $\mathsf{Ru} \subseteq \{R1, R2, R3, R4, R7\}$*, then L is hyperformal.*

Corollary 7. *DR^- is hyperformal.*

Proof. By inspection, DR^-'s axioms are closed under depth substitutions, so the previous corollary gives the result. □

Say that a sub*system* of a logic L is a logic generated from some subset of L's axioms and rules. We then have

Theorem 8. *If L is a subsystem of DR^-, then L is hyperformal.*

6. What's Wrong With DR?

In the original version of this paper, I claimed the above results held for all of DR. Tore Fjetland Øgaard pointed out in correspondence that this is incorrect. The problems stem from the disjunctive rules R4, R5, and R6. The argument I give above does not generalize to the disjunctive rules R4-R6.

In a bit more detail, suppose we have a derivation D of length $n+1$ ending at B and the last step in D looks like this:

$$\frac{A \qquad A \to B}{B}$$

Choose a depth substitution d. By the inductive hypothesis both $A[d]$ and $(A \to B)[d-1]$ are in DR. But then so are both $A[f]$ and $A[d] \to B[d]$. Thus since DR is closed under R2, $B[d]$ is in DR.

The same argument does not work in the case of R4: $(C \vee A)[d] = C[d] \vee A[d]$, but $(C \vee (A \to B))[d-1] = C[d-1] \vee (A[d] \to B[d])$. But the first disjuncts of the resulting formulas don't match so R4 no longer applies.

In the case of R5, $(C \vee \neg(A \to \neg A))[d] = C[d] \vee \neg(A[d+1] \to \neg A[d+1])$. But from $C \vee A$ being provable we can only guarantee that $(C \vee A)[d+1] = C[d+1] \vee A[d+1]$ and $(C \vee A)[d] = C[d] \vee A[d]$ are provable. The first of these shows that $C[d+1] \vee \neg(A[d+1] \to \neg A[d+1])$ is provable; the second that $C[d] \vee \neg(A[d] \to \neg A[d])$ is provable. Neither suffices for showing that $(C \vee \neg(A \to \neg A))[d]$ is provable. An exactly parallel observation shows a problem with the R6 clause in the induction.

Indeed, as Øgaard observed, not only does my argument not work, there are in fact counterexamples to the results stated in the original paper. As an example, $\neg p \vee \neg(p \to \neg p)$ is a theorem of DR but $\neg q \vee \neg(p \to \neg p)$, which is clearly a depth-substitution instance of it, is not. Thus the claim that DR is closed under depth-substitutions is false.

Finally, Øgaard also points out that the fact that DR has the depth-relevance property but is not hyperformal is itself a useful observation, as it demonstrates that hyperformality is a strictly stronger condition than depth-relevance. I suspect (but as of this writing have not proved) that in fact DR has the strong depth-relevance property. If so, DR would also serve to demonstrate that hyperformality is a strictly stronger condition than strong depth-relevance.

7. Connecting the Pieces

In the introduction we said that we'd give simplified proofs of results that subsumed Brady's results for DR$^-$. It's time to make good on this promise.

The basic idea is simple enough: the fact that DR$^-$ is hyperformal lets us transform any purported violation of the Strong Brady Theorem in DR into a violation of the Strong Belnap Theorem in DR and any purported violation of the Weak Brady Theorem in DR$^-$ into a violation of the Weak Belnap theorem in DR$^-$. But DR$^-$ is a sublogic of R. So by the Belnap Theorems there can be no such violations. QED.

The proof we give will give below is actually a proof of the contrapositive result. Nonetheless, the idea that Brady-counterexamples give rise to Belnap-counterexamples is at its heart.

In spite of how long you've been waiting, I won't dive straight into the proof here. This is because the proof is actually quite straightforward, *provided* the idea of it is completely clear. So it's useful to spend a moment working through some nice, concrete examples to see the idea in action.

To help in the discussion, we will first define the following phrases:

- We say that $A \to B$ *has the variable sharing property* when some variable p occurs in both A and B.

- We say that $A \to B$ *has the strong variable sharing property* when some variable p occurs in both A and B with the same sign
- We say that $A \to B$ *has the depth relevance property* when some variable p occurs at the same depth in both A and B.
- We say that $A \to B$ *has the strong depth relevance property* when some variable p occurs at the same depth in both A and B with the same sign.

Thus the Weak Belnap Theorem amounts to the claim that every provable conditional of every sublogic of R has the variable sharing property; the Strong Belnap Theorem to the claim that every provable conditional of every sublogic of R has the strong variable sharing property; the Weak Brady Theorem to the claim that every provable conditional in every sublogic of DR has the depth relevance property; and the Strong Brady Theorem to the claim that every provable conditional in every sublogic of DR has the strong depth relevance property.

It will also help our discussion to settle on a particular *injective* (i.e. one-one) depth substitution whose range includes only variables. For this purpose, we will use the function d that is defined as follows:

	p_1	p_2	p_3	p_4	p_5	p_6	$\cdots$
0	p_1	p_3	p_6	p_{10}	p_{15}	p_{21}	$\cdots$
1	p_2	p_5	p_9	p_{14}	p_{20}	$\cdots$	
2	p_4	p_8	p_{13}	p_{19}	$\cdots$		
3	p_7	p_{12}	p_{18}	$\cdots$			
4	p_{11}	p_{17}	$\cdots$				
5	p_{16}	$\cdots$					
$\vdots$	$\vdots$						

The claim I want to defend is that whenever $A \to B$ lacks the depth relevance property, $(A \to B)[d]$ will lack the variable sharing property and that whenever $A \to B$ lacks the strong depth relevance property, $(A \to B)[d]$ will lack the strong variable sharing property. As mentioned, it's useful to first look at some examples. But first a note: unlike in our discussion of the Belnap Theorems, when we apply a depth substitution to a conditional formula, we *do not* ignore the arrow in the middle. Thus, as a rather elementary example, while $p_1[d] = p_1$, $(p_1 \to p_1)[d] = p_2 \to p_2$.

Example 4: Consider first the bane of the Strong Belnap Theorem: $p_1 \to (p_2 \to p_1)$. This formula *has* the strong variable sharing property, since p_1 occurs positively in both its antecedent and its consequent. But it lacks the depth relevance property, since p_1 is a depth 0 subformula of the antecedent but not of the consequent.

Now let's apply d to it:

$$
\begin{aligned}
(p_1 \to (p_2 \to p_1))[d] &= d(p_1, 1) \to (d(p_2, 2) \to d(p_1, 2)) \\
&= p_2 \to (p_8 \to p_4)
\end{aligned}
$$

As promised, this formula lacks the variable sharing property. End Example

Example 5: Now consider (a version of) the formula we used to show that the Strong Brady Theorem was stronger than the conjunction of the Weak Brady Theorem and the Strong Belnap Theorem:

$$(p_1 \wedge (p_2 \to p_3)) \to ((p_3 \to p_4) \vee (p_4 \to p_1))$$

As we mentioned in Example 3, this formula has both the strong variable sharing property and the depth relevance property. But it lacks the strong depth relevance property. While I'll leave it to the reader to check my work, when we apply d to this formula the result is the following:

$$(p_2 \wedge (p_8 \to p_{13})) \to ((p_{13} \to p_{19}) \vee (p_{19} \to p_4))$$

This formula *does* have the variable sharing property because p_{13} occurs in both the antecedent and the consequent. But it does not have the strong variable sharing property—the instance of p_{13} in the antecedent occurs positively as a subformula of the antecedent while the instance in the consequent is negative as a subformula of the consequent. Thus, as claimed, applying d to a formula that lacked the strong depth relevance property resulted in a formula that lacked the strong variable sharing property. End Example

In order to turn these examples into an actual proof, we need to make a few observations. First, notice that applying a depth substitution to a formula leaves the *skeleton* of that formula—roughly speaking, the arrangement of connectives and parentheses in the formulas—intact. Second, note that when I apply d, each instance of a variable in the image is the image of an instance of a variable with the same sign in the original formula. For example, looking at example 4, notice that the instance of p_8 in the image of $p_1 \to (p_2 \to p_1)$ under d—that is to say, in $p_2 \to (p_8 \to p_4)$—is the image of p_2. And, just as p_8 occurs negatively in $p_2 \to (p_8 \to p_4)$, so also p_2 occurs negatively in $p_1 \to (p_2 \to p_1)$. It's not hard to see that in fact this is always the case. These observations will suffice for the proof.

Theorem 9. *Provable conditionals in hyperformal sublogics of R have the strong depth relevance property.*

Proof. Suppose L is a hyperformal sublogic of R and $\vdash_L A \to B$. Consider $(A \to B)[d]$. By our above observations, for some C and D, $(A \to B)[d]$ will have the form $C \to D$. Since L is hyperformal, $\vdash_L (A \to B)[d]$, which is to say that $\vdash_L C \to D$. Since L is a sublogic of R, it follows from the Strong Belnap Theorem that some variable q occurs with the same sign σ in both C and D.

Since q occurs in C with sign σ, there is a variable r_A that occurs at depth n_A with sign σ in A such that $d(r_A, n_A + 1) = q$.[3] Similarly, since q occurs in D with sign σ, there is a variable r_B that occurs at depth n_B with sign σ in B such that $d(r_B, n_B + 1) = q$. But then $d(r_A, n_A + 1) = q = d(r_B, n_B + 1)$. So by the injectivity of d, $r_A = r_B$ and $n_A = n_B$. Thus some variable $r = r_A = r_B$ occurs with sign σ at the same depth, $n = n_A = n_B$ in both A and B. So $A \to B$ has the strong depth relevance property. □

As a corollary we have

Corollary 10 (The Strong Brady Theorem)**.** *If L is a sublogic of the logic DR^- and $\vdash_L A \to B$, then some variable p occurs at the same depth and with the same sign in both A and B.*

Proof. Clearly if L is a sublogic of DR^- and $\vdash_L A \to B$, then $\vdash_{DR^-} A \to B$ as well. By Corollary 7, DR^- is hyperformal. So by Theorem 9, if $\vdash_{DR^-} A \to B$, then $A \to B$ has the strong depth relevance property, which was what we wanted to show. □

Finally, note that we managed this result without any of the odd machinery required in the proof of the Brady Theorems. I take this to be an advantage.

[3]Since r_A occurs at depth n_A in A it occurs at depth $n_A + 1$ in $A \to B$.

8. Conclusion

The Belnap Theorems say something about relevance qua variable sharing. So do the Brady Theorems. The Brady Theorems are stated for sublogics of DR^-. But the sublogics of DR^- we are mostly interested in are all subsystems of DR^-. And it turns out that subsystems of DR^- are hyperformal. In the presence of hyperformality, the Brady Theorems reduce to the Belnap Theorems in the sense that any violation of the former can be turned into a violation of the latter. It follows that, for such logics, the Brady Theorems are in a sense the natural hyperformal analogue of the Belnap Theorems.

What remains is to plumb the philosophical significance of hyperformality. Prior to stating and proving the Strong Belnap Theorem in [1], Anderson and Belnap say that it shows that 'if $A \to B$ is provable. . . then A and B share *intensional content*, in the sense that they share a variable.' (emphasis added). Presumably the Brady Theorems thus demonstrate that A and B share something stronger than whatever 'intensional content' means to Anderson and Belnap. It seems that hyperformality could shed light on what, exactly this strong shared-something is. I haven't yet thought hard enough about it to say whether this intuition is correct. But DR^- is fairly close to Brady's logic of meaning containment MC (see [4]) which is in turn explicitly designed to capture the notion of intensional content. Thus, a plausible way we might shed light on this subject is by examining MC.

Acknowledgments This paper benefited from the audiences of talks on various parts of this material at the Nonclassical Logic Webinar and at the Annual Meeting of the Australasian Association for Logic. In particular, I thank Graham Priest, Ross Brady, Dave Ripley, Ed Mares, Sara Ugolini, Shawn Standefer, and Guillermo Badia. Like much of my work of late, this paper also owes a great deal to the hive mind comprised of Andrew Tedder, Graham Leach-Krouse, Eileen Nutting, and Teresa Kouri. Finally, two referees for this journal were instrumental in my catching several crucial errors, for which I am quite grateful.

References

[1] Alan Ross Anderson and Nuel D. Belnap. *Entailment: The Logic of Relevance and Necessity*. Princeton University Press, 1975.
[2] Nuel D Belnap. Entailment and relevance. *The Journal of Symbolic Logic*, 25(2):144–146, 1960.
[3] Ross T. Brady. Depth relevance of some paraconsistent logics. *Studia Logica*, 43(1-2):63–73, 1984.
[4] Ross T. Brady. *Universal Logic*. CSLI Publications, 2006.
[5] Shay Allen Logan. Strong depth relevance. *Australasian Journal of Logic*, 18(6):645–656, 2021.
[6] John MacFarlane. *What Does It Mean to Say That Logic is Formal?* PhD thesis, University of Pittsburgh, 2000.