Proof Invariance

Blane Worley

University of California, Davis

Abstract

We explore depth substitution invariance, or hyperformalism, and extend known results in this realm to justification logics extending weak relevant logics. We then examine the surprising invariance of justifications over formulas and restrict our attention to the substitution of proofs in the original relevant logic. The results of this paper indicate that depth invariance is a recalcitrant feature of the logic and that proof structures in hyperformal logics are quite inflexible.

1 Introduction

Relevance is a notoriously tricky property to pin down. Early attempts to formalize the notion of relevance in propositional logic focused on shared variables between premise and consequent (Belnap (1960), Anderson and Belnap (1975)). A logic **L** is said to enjoy the *variable sharing property* (vsp) if whenever $A \to B$ is a theorem of **L**, then A and B share an atom. vsp is an oft-cited necessary condition for relevance, but it is not a sufficient condition.[1] Furthermore, as shown in Anderson and Belnap (1975) and thoroughly explored in Méndez and Robles (2012), this property can be significantly strengthened (svsp, for *strong* vsp) over the domain of standard relevant logics which already enjoy vsp, though one can construct (irrelevant) logics which enjoy the standard vsp but not the svsp.

A yet stronger property, known as the depth relevance property (see Brady (1984)), has also been extensively explored. In fact, Robles and Méndez (2014) suggest that depth relevance is a potentially "fitter condition than the vsp to characterize relevant logics" (p. 125). Depth relevance can be strengthened in a similar way to how the vsp was strengthened (Logan (2021)), but we will be concerned in this paper primarily with the original formulation. Formulas A and B *depth share a variable* if some atomic variable occurs in both A and B at the same depth.

Definition 1. A logic **L** enjoys the *depth relevance property* (drp) if $A \to B$ is a theorem of **L** only if A and B depth share a variable.

[1] The fragment of classical logic that enjoys vsp, for example, admits disjunctive syllogism, which is a rejected principle of relevance. See Szmuc (2021) for a sequent calculus and semantics for this fragment, with which it is rather simple to show the validity of the offending principle. See Mares (2004) for a salient examination of this deductive principle in the context of inconsistency.

A formal definition of depth is given further in the paper. Informally, the depth of a subformula in a formula is the number of conditionals under which the subformula is nested in the formula.

Strong relevant logics do not satisfy this condition. **R** is a fine counterexample, since it contains instances of *assertion*: $p \to ((p \to q) \to q)$ for example. The failure here can be seen by noting that in the antecedent p, the variable p occurs at depth 0, since it occurs within the scope of no other conditional, while the p in the consequent occurs at depth 2. Since p was the only contender for a shared atom, this is enough to be a counterexample. drp is thus applicable to a smaller subclass of logics under **R** than is vsp.[2]

In the course of building machinery to identify the logics that enjoy drp, Logan (2022) employs substitution-like mappings drawn from a class of functions he calls *depth substitutions*. Like uniform substitutions, depth substitutions map arbitrary atomic variables to sentences. Unlike uniform substitutions, depth substitutions need not treat all occurrences of a given variable equally. Instead, depth substitutions are functions which take each instance of a variable at a given depth to a specified formula. In some logics this arbitrary mapping does not affect the space of theorems.

Definition 2. A logic is said to be *hyperformal* whenever it admits a theorem just if it also admits every depth substitution of that theorem.

A sufficient condition for depth relevance is then established.

Theorem 1 (Logan (2022))**.** *If a logic is hyperformal, then it also enjoys the depth relevance property.*

While the converse of this theorem is not true, there is a large overlap between depth relevant logics and hyperformal logics. It was shown in Brady (1984) that sublogics of **DR** satisfy the drp and in Logan (2022) that sublogics of $\mathbf{DR}^-$ are hyperformal. It is clear then that for weak-enough relevant logics, these notions coincide.

Depth substitutions are indeed strange. Nonuniform mappings would not typically be expected to preserve theoremhood. The classical logician would accept $p \to (q \to p)$ among their available tautologies but would obviously reject its depth substitution instance $p \to (q \to r)$. In particular, we can construct a depth substitution which sends

$$p \to (q \to p) \rightsquigarrow (p \to p) \to ((q \to q) \to r)$$

and with a small basis of logical power, namely identity and modus ponens, one gets that r is a theorem. Since r was atomic and a logic should be closed under uniform substitution, it follows that all formulas are theorems.

How then should we go about those nontrivial logics that *are* closed under such an unusual class of functions? It's not obvious at this juncture whether this feature is worth taking seriously, since

[2] If taken as a necessary criteria for relevance, drp is also not a sufficient condition. An example from Robles and Méndez (2014) gives a category of **RM3** extensions which all respect drp while admitting principles incompatible with relevance. An illustrative axiom scheme one could include among the theorems of, say, **DR** is the "intractable principle" $(\phi \to \psi) \to (\psi \to \psi)$.

being closed under a property does not necessarily carry any philosophical weight. For example, relevant logics are closed under the rule $A \wedge \neg A \vdash B$, but this observation doesn't convey much information about the logics themselves. We must then justify, so to speak, that closure under depth substitutions is a feature worth us causing a scene.

The purpose of the present paper is to show that there is, in fact, good reason to cause a scene. We do this by showing hyperformalism is embedded not simply into the theorems of hyperformal logics but (in a sense made clear below) in the proofs of those theorems as well. To do so, we employ the tools of justification logic and use the work done by Standefer (2022) on variable sharing with justification to derive related results about depth relevance. Justification logics, much like modal logics, augment a propositional logic with structure tracking whichever modal behavior one is concerned with, *e.g.* belief, knowledge, obligation, and provability. Informally, justification logics "represent the whys and wherefores of modal behavior quite directly, and from within the formal language itself" (Artemov and Fitting, xii). We will concern ourselves with the notion of proof and provability and will thus read formulas involving justifications as saying something about the provability of the formulas to which they're attached.

In §2, the logics **B** and **B**.J_0 are defined, and details surrounding how each justification is meant to capture provability are discussed briefly. In §3, we define depth and introduce depth substitutions. After that, constant specifications are defined and philosophically motivated in §4, followed in §5 by the substantive content of the paper in which we show that, in hyperformal logics, proof terms of a theorem are more or less attached to every depth substitution instance of that theorem. Finally, §6 looks at a subset of the justification logic to motivate a natural perspective for this behavior in light of the extant literature on depth substitutions, and we summarize our results in the final remarks.

2 The Logic

We begin with a set JV of enumerably many justification variables, a yet unspecified (and possibly empty) set JC of justification constants, and a set of connectives $\mathsf{Con} = \{+, \cdot, \mathfrak{a}, \mathfrak{b}, \mathfrak{c}, *, !, \neg, \wedge, \vee, \rightarrow\}$. We form the set JT of *justification terms* inductively.

- If $x \in \mathsf{JV}$, then $x \in \mathsf{JT}$
- If $c \in \mathsf{JC}$, then $c \in \mathsf{JT}$
- Given $t \in \mathsf{JT}$ and $s \in \mathsf{JT}$,
  - $(t + s) \in \mathsf{JT}$
  - $(t \cdot s) \in \mathsf{JT}$
  - $\mathfrak{a}(t, s) \in \mathsf{JT}$
  - $\mathfrak{b}(t, s) \in \mathsf{JT}$
  - $\mathfrak{c}t \in \mathsf{JT}$
  - $*t \in \mathsf{JT}$

– $!t \in \mathsf{JT}$

Now assume we have an enumerable set At of propositional constants. We form *justification formulas* as follows.

- If $p \in \mathsf{At}$, p is a justification formula.
- If ϕ and ψ are justification formulas, so then are $\neg\phi$, $\phi \wedge \psi$, $\phi \vee \psi$, and $\phi \rightarrow \psi$.
- If ϕ is a justification formula, then $[t]:(\phi)$ is a justification formula for all $t \in \mathsf{JT}$.

We will omit the brackets and parentheses in justification formulas whenever confusion should not arise, as in $\mathfrak{a}(t,s):(\phi \wedge \psi)$ and $s:t:\phi$. We call the language so-defined $\mathbf{L}^+$, and we denote the set of all formulas of the form $t:\phi$, where t may be an empty placeholder, by JF. The propositional fragment $\mathcal{L}$ of $\mathcal{L}^+$ is obtained by isolating the set $\mathsf{WFF} \subset \mathsf{JF}$ of justification-less formulas and the subcollection $\{\neg, \wedge, \vee, \rightarrow\} \subset \mathsf{Con}$. The logic **B** is presented in the language $\mathcal{L}$ with a minimal Hilbert-style axiomatization and is taken as the smallest set that contains every instance of the following axioms and is closed under the following inference rules.

(A1) $\phi \rightarrow \phi$

(A2) $(\phi \wedge \psi) \rightarrow \phi, (\phi \wedge \psi) \rightarrow \psi$

(A3) $((\phi \rightarrow \psi) \wedge (\phi \rightarrow \lambda)) \rightarrow (\phi \rightarrow (\psi \wedge \lambda))$

(A4) $\phi \rightarrow (\phi \vee \psi), \psi \rightarrow (\phi \vee \psi)$

(A5) $((\phi \rightarrow \lambda) \wedge (\psi \rightarrow \lambda)) \rightarrow ((\phi \vee \psi) \rightarrow \lambda)$

(A6) $(\phi \wedge (\psi \vee \lambda)) \rightarrow ((\phi \wedge \psi) \vee (\phi \wedge \lambda))$

(A7) $\neg\neg\phi \rightarrow \phi$

(R1) $\dfrac{\phi \rightarrow \psi \qquad \phi}{\psi}$

(R2) $\dfrac{\phi \qquad \psi}{\phi \wedge \psi}$

(R3) $\dfrac{\phi \rightarrow \neg\psi}{\psi \rightarrow \neg\phi}$

(R4) $\dfrac{\phi \rightarrow \psi \qquad \lambda \rightarrow \rho}{(\psi \rightarrow \lambda) \rightarrow (\phi \rightarrow \rho)}$

We formulate the logic $\mathbf{B}.\mathrm{J}_0$ over $\mathcal{L}^+$ by including among (A1)-(A7) and (R1)-(R4) the following axioms and one additional rule.

(J0) $t : \phi \rightarrow [t+s]:\phi,\ s : \phi \rightarrow [t+s]:\phi$

(J1) $t : (\phi \rightarrow \psi) \rightarrow (s : \phi \rightarrow [t \cdot s] : \psi)$

(J2) $(t : \phi) \wedge (s : \psi) \rightarrow \mathfrak{a}(t,s) : (\phi \wedge \psi)$

(J3) $t : (\phi \rightarrow \neg\psi) \rightarrow \mathfrak{c}t : (\psi \rightarrow \neg\phi)$

(J4) $t:\phi \rightarrow (!t:t:\phi)$

(JR5) $\dfrac{t:(\phi \rightarrow \psi) \qquad s:(\lambda \rightarrow \rho)}{\mathfrak{b}(t,s):((\psi \rightarrow \lambda) \rightarrow (\phi \rightarrow \rho))}$

If $t:\phi$ is a theorem, we will sometimes refer to t as a *proof term* for ϕ. (J0) and (J1) provide access to the basic operators $+$ and $\cdot$ among justification terms, which is why these terms are sometimes called justification polynomials. When taken in isolation, (J1) and (J4) form an epistemic analogue of **K4**, which includes as axioms both $K(\phi \rightarrow \psi) \rightarrow (K\phi \rightarrow K\psi)$ and $K\phi \rightarrow KK\phi$. The remaining axioms and rule involving explicit justification terms correspond to more trivially satisfied modal statements or else to a rule in **B**. The other justification operators ($\mathfrak{a}$, $\mathfrak{b}$, $\mathfrak{c}$, and !) combine justification terms, and they intuitively correspond to different ways of combining actual justifications. Note that (JR5), which corresponds to (R4), is given as an additional inference rule instead. This will cause us a minor headache in §5, but a brief discussion of the rationale for this decision will follow there.

The main result will be proved for $\mathbf{L}.\mathrm{J}_0$ where $\mathbf{L} = \mathbf{B}$, but we note here that this result will apply to a broad enough range of relevant logics (see Lemma 6 below). Somewhere around the supremum of this range is the logic $\mathbf{DR}^-$. Since we will make repeated mentions of this logic throughout the paper, an explicit axiomatization is laid out here. The logic $\mathbf{DR}^-$ is the logic **B**, minimally closed under the below additional axioms and rule.

(A8) $((\phi \to \psi) \land (\psi \to \lambda)) \to (\phi \to \lambda)$

(A10) $\phi \lor \neg\phi$

(A9) $(\phi \to \neg\psi) \to (\psi \to \neg\phi)$

(R5) $$\frac{\phi}{\neg(\phi \to \neg\phi)}$$

An informed reader will notice how closely this logic resembles the well-studied and philosophically important logic **DJ**, and indeed **DJ** is a sublogic of $\mathbf{DR}^-$. Of course, we will need to add an additional justification axiom or rule to account for uses of (R5). To achieve this goal, we will add the following rule to our system $\mathbf{DR}^-.\mathrm{J}_0$.

(JR6) $$\frac{t : \phi}{*t : \neg(\phi \to \neg\phi)}$$

Here, $*$ encodes uses of (R5) in much the same way as the other justification connectives encode their corresponding rules. In general, however, most extensions of **B** we will consider involve *fewer* rules, rather than additional rules. To avoid unnecessary additional rules, we will consider a restricted class of sublogics of $\mathbf{DR}^-$, here called subsystems, for the remainder of this paper.

Definition 3. A *subsystem* of a logic **L** is a sublogic of **L** whose axiomatization includes only rules from (R1)-(R5).

A final desideratum for this new structure would be something analogous to necessitation in normal modal logic, that if ϕ is a theorem, then so too is $K\phi$. There is an analogous class of results in the literature called *internalization* theorems. These take the form "if ϕ is a theorem, there is a t so that $t : \phi$ is a theorem." Internalization is a desirable feature for our logic to have. Standefer (2022) demonstrates that his formulation of a justification logic over **B**, under certain further constraints (see §4), satisfies internalization. We will rederive this result in our formulation and in a separate context.

3 Depth

The notion of depth underlying the depth relevance principle (see Brady (1984)) is extended in the following for formulas in the full language $\mathcal{L}^+$.[3] We will first state the standard definition for depth and depth substitution over the propositional fragment $\mathcal{L}$ of our language. After this, we will consider how to extend this definition to cover the larger language $\mathcal{L}^+$.

Definition 4. Let $\phi \in \mathcal{L}$ and $\psi \in \mathcal{L}$. Denote by ψ^ϕ a particular occurrence of ϕ in ψ. The *depth of an occurrence of ϕ in ψ* is a non-negative integer given inductively by the following.

- If ϕ is ψ, the depth of ψ^ϕ is 0.
- If the depth of ψ^ϕ is n, then
  - The depth of $(\neg\psi)^\phi$ corresponding to ψ^ϕ is n.
  - The depth of $(\psi \vee \lambda)^\phi$ and of $(\lambda \vee \psi)^\phi$ corresponding to ψ^ϕ is n.
  - The depth of $(\psi \wedge \lambda)^\phi$ and of $(\lambda \wedge \psi)^\phi$ corresponding to ψ^ϕ is n.
  - The depth of $(\psi \to \lambda)^\phi$ and of $(\lambda \to \psi)^\phi$ corresponding to ψ^ϕ is $n + 1$.

Recall that in the construction of our language $\mathcal{L}^+$, we specified a set of atomic constants At.

Definition 5. A *depth substitution* is a function $d^- : \mathsf{At} \times \mathbb{Z} \to \mathcal{L}$.[4] This extends to a function (which we will, as an abuse of terminology, also call a depth substitution) $d : \mathcal{L} \times \mathbb{Z} \to \mathcal{L}$ as follows:

- $d(p, n) = d^-(p, n)$ for $p \in \mathsf{At}$
- $d(\neg\phi, n) = \neg d(\phi, n)$
- $d(\phi \vee \psi, n) = d(\phi, n) \vee d(\psi, n)$
- $d(\phi \wedge \psi, n) = d(\phi, n) \wedge d(\psi, n)$
- $d(\phi \to \psi, n) = d(\phi, n + 1) \to d(\psi, n + 1)$

Whenever we refer to an arbitrary depth substitution d, the extension is our intention. A *uniform* substitution is a degenerate depth substitution, one in which $d(\phi, n) = d(\phi, 0)$ for all $n \in \mathbb{Z}$. An interesting result alluded to earlier is due to Logan (2022).

Theorem 2. *Let* $\mathbf{L}$ *be a sublogic of* $\mathbf{DR}^-$ *and* $\phi \in \mathbf{L}$*. Given a depth substitution* d *and* $n \in \mathbb{Z}$*,* $d(\phi, n) \in \mathbf{L}$.

[3] Anderson and Belnap (Anderson and Belnap (1975)) conceptualize an early consideration of depth in terms of a formula's *degree* in the study of their token relevant logics. The standard formulation, however, is due to Brady (1984).

[4] In Logan (2022), depth spans over $\mathbb{N}$ since negative depth does not occur when considering propositional logics of this flavor. However, we need to have a contingency for negative depth due to the depth-shifting nature of (J1) and (JR5), even if we don't know what it means for ϕ to occur at depth -3 in ψ. The philosophical significance of negative depth will not be explored here.

How then should a depth substitution treat justification terms in these weak logics, in light of Theorem 2? We argue that depth and depth substitutions should be blind about justification terms, just as they are blind to formulas combined with other non-conditional connectives. That is, we extend Definitions 4 and 5 to range over formulas of $\mathcal{L}^+$ with the following respective conditions.

- The depth of $(t:\psi)^\phi$ corresponding to ψ^ϕ is n.
- For a depth substitution $d\colon \mathcal{L}^+ \times \mathbb{Z} \to \mathcal{L}^+$, $d(t:\phi, n) = t:d(\phi, n)$

Arbitrary depth substitutions through the remainder of this paper will be named instances of this final extension. The effect of this imposition is that proofs of a formula are also proofs of their depth substitutions. We must take a small measure of this epistemological choice, as it may appear arbitrary on first viewing. Depth is, as defined, a property tracking the number of occurrences of one particular connective. The term *depth* is itself suggestive of which connective we're interested in tracking, but we could have just as easily defined the behavior of our substitutions so that they only track the number of occurrences of negations, conjunctions, or disjunctions under which a formula is nested within a broader formula.[5] The inert nature of depth and depth substitutions over all other connectives is itself an authorial choice, whereas the fact that this choice aligns with quite interesting results, such as Theorem 2 above, is a consequence of this choice.

Justification terms, however, stand in some middle ground between object language connectives and metainferential objects. Theorems containing justification terms are in essence object-level representations of our logical rules, telling us precisely which, and in what order, rules were employed in deriving the attached formula, and can be deconstructed, much like parsing a proof tree. Further, justification terms have intensional content—in the formal language, each justification term gives rise to a connective that is just as intensional as a conditional. It is thus reasonable to request that proof terms have some non-trivial effect on a formula's depth. Why then ignore this intensionality? Well, the choice to ignore alternative intensional structure was already made when we decided to ignore the intensional nature of negations. We have thus chosen to treat attached justification terms as any other non-conditional connective. In doing so, we acknowledge that alternative interpretations of depth may credibly be given but find our interpretation most compelling in light of the above.

A hyperformal logic then, in accordance with Theorem 1, should be expected to fail to distinguish between proofs of statements of the same form, since it fails to distinguish between the statements already. Thus we'd expect a similar characteristic for justification logics, but we need just a little bit more tooling to see whether this is the case.

[5]We could even redefine our substitutions and tweak the underlying notion (here, depth) so that both track the use of multiple connectives and the order in which they occur. Alas, these substitutions have just in the last year been examined and probed in earnest, as in the forthcoming Ferguson and Logan, with promising early results. The current paper restricts its attention to depth substitutions, though a note on extending to broader classes of substitutions will be given in Section 7.

4 Constant Specifications

So far, the justificational superstructure we've adjoined to **B** does not generate genuine justificational theorems, i.e. there is no theorem in $\mathbf{B}.\mathrm{J}_0$ that is of the form $t\!:\!\phi$. The subscripted "0" in the logic's name is used to indicate precisely this fact. We can ensure theorems of this form by specifying the set JC in $\mathcal{L}^+$ and indicating to which theorems of $\mathbf{B}.\mathrm{J}_0$ these atomic proof terms will attach. The resulting set is called a constant specification and will be treated carefully throughout the remainder of this paper.

Definition 6. A *constant specification* for a justification logic $\mathbf{L}.\mathrm{J}_0$ is a set $\mathsf{CS} \subseteq \mathsf{JF}$ that satisfies the following.

- Elements of CS are of the form $c_n\!:\!c_{n-1}\!:\ldots:\!c_1\!:\!\phi$, where c_i is a justification constant for $0 < i \leq n$ and ϕ is an *axiom* of $\mathbf{L}.\mathrm{J}_0$.
- If $c_n\!:\!c_{n-1}\!:\ldots:\!c_1\!:\!\phi$ is in CS where $n > 1$, then so is $c_{n-1}\!:\ldots:\!c_1\!:\!\phi$.

If we add the members of a constant specification CS to a justification logic $\mathbf{L}.\mathrm{J}_0$ as axioms, the resulting system is usually labelled $\mathbf{L}.\mathrm{J}_{\mathsf{CS}}$, and we adopt this notation in this paper. There are many flavors of constant specifications, and these can be tweaked to fit the logician's purpose. A few candidates are outlined here.

Schematic If ϕ and ψ are both instances of the same axiom scheme, $c\!:\phi \in \mathsf{CS}$ if and only if $c\!:\!\psi \in \mathsf{CS}$, for every justification constant c.

Axiomatically appropriate (AACS) If ϕ is an instance of an axiom and $n > 0$, there are constants $c_i \in \mathsf{JC}$ for $1 \leq i \leq n$ so that $c_n\!:\ldots:\!c_1\!:\!\phi \in \mathsf{CS}$.

As mentioned in §2, a first desideratum for a plausible constant specification CS over a given justification logic $\mathbf{L}.\mathrm{J}_0$ is that the logic $\mathbf{L}.\mathrm{J}_{\mathsf{CS}}$ should satisfy internalization. As we will see, an AACS easily grants us this criterion. A second wish is that justifications for theorems in the base logic (what will later be called *first-degree proof terms*) somehow mirror the structure of derivations of the theorem. That is, we wish for single-length proof terms t of theorems $\phi \in \mathbf{B}$ to correspond bijectively with proofs of ϕ.

We could specify a specialized constant specification that seems to meet both these criteria, but as we will see, a surprising result is that this feature follows simply from internalization, the definition of depth substitution, and the hyperformal characteristic of the given base logic. Thus, given the above comment, we will explore an axiomatically appropriate constant specification for the time being.

5 Hyperformal Justification Logics

Now we show that there is a deeply encoded invariance of justifications for theorems of $\mathbf{B}.\mathrm{J}_{\mathsf{CS}}$, provided CS is axiomatically appropriate. To see this, however, we need to first establish a similar result to Theorem 2 and demonstrate internalization.

Lemma 3. *Given any constant specification* CS*, if* $\phi \in \mathbf{B}.\mathrm{J}_{\mathsf{CS}}$*, then for any depth substitution* d *and any* $n \in \mathbb{Z}$*,* $d(\phi, n) \in \mathbf{B}.\mathrm{J}_{\mathsf{CS}}$.

Proof. Let d be a depth substitution and $n \in \mathbb{Z}$. We proceed by induction on complexity of the derivation of ϕ. For axioms and rules of **B**, we capitulate to Theorem 2, which covers precisely this case. For the remaining axioms and rule, we argue as follows.

(J1) ϕ is of the form $t\!:\!(\phi_1 \to \phi_2) \to (s\!:\!\phi_1 \to [t \cdot s]\!:\!\phi_2)$. Then

$$\begin{aligned} d(\phi, n) &= d(t\!:\!(\phi_1 \to \phi_2) \to (s\!:\!\phi_1 \to [t \cdot s]\!:\!\phi_2), n) \\ &= d(t\!:\!(\phi_1 \to \phi_2), n+1) \to d(s\!:\!\phi_1 \to [t \cdot s]\!:\!\phi_2, n+1) \\ &= t\!:\!d(\phi_1 \to \phi_2, n+1) \to (d(s\!:\!\phi_1, n+2) \to d([t \cdot s]\!:\!\phi_2, n+2)) \\ &= t\!:\!(d(\phi_1, n+2) \to d(\phi_2, n+2)) \to (s\!:\!d(\phi_1, n+2) \to [t \cdot s]\!:\!d(\phi_2, n+2)) \end{aligned}$$

which is again an instance of axiom (J1).

(J2) ϕ is of the form $(t\!:\!\phi_1 \wedge s\!:\!\phi_2) \to \mathfrak{a}(t, s)\!:\!(\phi_1 \wedge \phi_2)$.

$$\begin{aligned} d(\phi, n) &= d((t\!:\!\phi_1 \wedge s\!:\!\phi_2) \to \mathfrak{a}(t, s)\!:\!(\phi_1 \wedge \phi_2), n) \\ &= d(t\!:\!\phi_1 \wedge s\!:\!\phi_2, n+1) \to d(\mathfrak{a}(t, s)\!:\!(\phi_1 \wedge \phi_2), n+1) \\ &= (d(t\!:\!\phi_1, n+1) \wedge d(s\!:\!\phi_2, n+1)) \to d(\mathfrak{a}(t, s)\!:\!(\phi_1 \wedge \phi_2), n+1) \\ &= (t\!:\!d(\phi_1, n+1) \wedge s\!:\!d(\phi_2, n+1)) \to \mathfrak{a}(t, s)\!:\!d(\phi_1 \wedge \phi_2, n+1) \\ &= (t\!:\!d(\phi_1, n+1) \wedge s\!:\!d(\phi_2, n+1)) \to \mathfrak{a}(t, s)\!:\!(d(\phi_1, n+1) \wedge d(\phi_2, n+1)) \end{aligned}$$

Thus $d(\phi, n)$ is an instance of axiom (J2).

The remaining axioms have similar arguments.

Now assume the lemma holds for derivations of length $m - 1$ and that ϕ is derived over m steps with the last step using (JR5). Then ϕ is of the form $\mathfrak{b}(t, s)\!:\!((\lambda_{12} \to \lambda_{21}) \to (\lambda_{11} \to \lambda_{22}))$, and the derivation of ϕ looks like this:

$$\frac{t\!:\!(\lambda_{11} \to \lambda_{12}) \qquad s\!:\!(\lambda_{21} \to \lambda_{22})}{\mathfrak{b}(t, s)\!:\!((\lambda_{12} \to \lambda_{21}) \to (\lambda_{11} \to \lambda_{22}))}$$

Note that

$$d(\phi, n) = \mathfrak{b}(t, s)\!:\!((d(\lambda_{12}, n+2) \to d(\lambda_{21}, n+2)) \to (d(\lambda_{11}, n+2) \to d(\lambda_{22}, n+2)))$$

and this is derivable by (JR5) from the following theorems.

$$\begin{aligned} &t\!:\!(d(\lambda_{11}, n+2) \to d(\lambda_{12}, n+2)), \text{ and} \\ &s\!:\!(d(\lambda_{21}, n+2) \to d(\lambda_{22}, n+2)) \end{aligned}$$

But by IH, we have that $d(t:(\lambda_{11} \to \lambda_{12}), n+1) \in \mathbf{B}.\mathsf{J}_{\mathsf{CS}}$ and $d(s:(\lambda_{21} \to \lambda_{22}), n+1) \in \mathbf{B}.\mathsf{J}_{\mathsf{CS}}$. These theorems are precisely the ones we need to prove the result. For instance,

$$\begin{aligned} d(t:(\lambda_{11} \to \lambda_{12}), n+1) &= t:d(\lambda_{11} \to \lambda_{12}, n+1) \\ &= t:(d(\lambda_{11}, n+2) \to d(\lambda_{12}, n+2)) \end{aligned}$$

and similarly for the s formula. Thus by (JR5), we have a derivation of $d(\phi, n)$. □

The next result is proved in Standefer (2022) but in a different presentation for **B**, in which additional inference rules and more connectives are present in the logic and language. We reprove it here and show the differing cases for the inductive step.

Lemma 4 (Internalization)**.** *Given an axiomatically appropriate constant specification* CS*, if* $\phi \in \mathbf{B}.\mathsf{J}_{\mathsf{CS}}$*, then there is a justification term* t *so that* $t:\phi \in \mathbf{B}.\mathsf{J}_{\mathsf{CS}}$.

Proof. Proceed by induction on the length of the derivation of ϕ. If ϕ is an instance of an axiom, then by our choice of constant specification there is a constant c so that $c:\phi \in \mathbf{B}.\mathsf{J}_{\mathsf{CS}}$. Suppose now that ϕ is derived over m inference steps and that the theorem is proved for theorems with derivation length $m-1$. A selection of cases is presented here; the remaining cases are left to the reader.

(R1) ϕ is derived as follows.

$$\frac{\psi \qquad \psi \to \phi}{\phi}$$

By IH, there are justification terms u, v so that $u:\psi$ and $v:\psi \to \phi$ are theorems. But then we can derive our result via R1 and J1.

$$\dfrac{u:\psi \qquad \dfrac{v:(\psi \to \phi) \qquad v:(\psi \to \phi) \to (u:\psi \to [u \cdot v]:\phi)}{u:\psi \to [u \cdot v]:\phi}}{[u \cdot v]:\phi}$$

Thus there is t—i.e. $[u \cdot v]$—so that $t:\phi \in \mathbf{B}.\mathsf{J}_{\mathsf{CS}}$.

(R4) ϕ is of the form $(\lambda_{12} \to \lambda_{21}) \to (\lambda_{11} \to \lambda_{22})$ with derivation

$$\frac{\lambda_{11} \to \lambda_{12} \qquad \lambda_{21} \to \lambda_{22}}{(\lambda_{12} \to \lambda_{21}) \to (\lambda_{11} \to \lambda_{22})}$$

By IH, there are u, v so that $u:(\lambda_{11} \to \lambda_{12}), v:(\lambda_{21} \to \lambda_{22}) \in \mathbf{B}.\mathsf{J}_{\mathsf{CS}}$. We receive our result immediately by JR5.

$$\frac{u:(\lambda_{11} \to \lambda_{12}) \qquad v:(\lambda_{21} \to \lambda_{22})}{\mathfrak{b}(u, v):((\lambda_{12} \to \lambda_{21}) \to (\lambda_{11} \to \lambda_{22}))}$$

(JR5) ϕ is of the form $\mathtt{b}(t_1, t_2) : ((\lambda_{12} \to \lambda_{21}) \to (\lambda_{11} \to \lambda_{22}))$ and is derived by the following.

$$\frac{t_1 : (\lambda_{11} \to \lambda_{12}) \qquad t_2 : (\lambda_{21} \to \lambda_{22})}{\mathtt{b}(t_1, t_2) : ((\lambda_{12} \to \lambda_{21}) \to (\lambda_{11} \to \lambda_{22}))}$$

Note that we do not need the full power of the inductive hypothesis here, since we included enough epistemic power in our justification axiom scheme to account for this scenario. Use (R1) and (J4) as appropriate, where $A = (\lambda_{12} \to \lambda_{21}) \to (\lambda_{11} \to \lambda_{22})$.

$$\frac{\mathtt{b}(t_1, t_2) : A \qquad \mathtt{b}(t_1, t_2) : A \to !\mathtt{b}(t_1, t_2) : \mathtt{b}(t_1, t_2) : A}{!\mathtt{b}(t_1, t_2) : \mathtt{b}(t_1, t_2) : A}$$

□

Of note is the last case considered in the above proof. What if we didn't add positive introspection into our system? Then the derivation seems out of reach and ultimately unsuccessful. One well-meaning attempt would have us produce a separate proof of the formula by introducing Fitting's $\mathtt{c}$ operator via the axiom schema $s : t : \phi \to [s \mathtt{c} t] : \phi$ (see Fitting (2017) on the logic **JX4**). So given the theorems (by IH) $u : t_1 : (\lambda_{11} \to \lambda_{12})$ and $v : t_2 : (\lambda_{21} \to \lambda_{22})$, we would have another derivation of the base formula:

$$\mathtt{b}([u \mathtt{c} t_1], [v \mathtt{c} t_2]) : ((\lambda_{12} \to \lambda_{21}) \to (\lambda_{11} \to \lambda_{22}))$$

But this isn't where we need to be! This and similar approaches using the tools at our disposal allow us only to manipulate the first proof term in a chain of possibly many, without looking a step beyond. Though we do not find this restriction a philosophical challenge as much as a formal constraint, an alternative presentation of the justification logic will grant us the tools needed to avoid this epistemic pigeonhole.

The case of the proof revolves around the rule (JR5). Altering the rule to be an analogous axiom would solve this problem, but we would have an immediate counterexample to drp, since any such axiom form one could propose does not adhere to this criteria, and thus a counterexample to hyperformalism.[6] One possible resolution is to remove (J4) and replace the single rule (JR5) with the schemes below, where (JR5) is now just the $n = 1$ case.

$$\text{(JRn)} \quad \frac{t_n : \ldots : t_1 : (\phi \to \psi) \qquad s_n : \ldots : s_1 : (\lambda \to \rho)}{b(t_n, s_n) : \ldots : b(t_1, s_1) : ((\psi \to \lambda) \to (\phi \to \rho))}, n \geq 1$$

With these new rule schemes, the induction goes through using the full power of the induction hypothesis and an additional induction on the length of the justification chain. For a given derivation by (JRn), appeal to (JR(n+1)) and IH in the proof. Whichever presentation of $\mathbf{B}.\mathsf{J}_{\mathsf{CS}}$ one adopts, internalization holds.

Finally, we make a simple but powerful observation.

[6]Take $t : (\phi \to \psi) \to (s : (\lambda \to \rho) \to \mathtt{b}(t, s) : ((\psi \to \lambda) \to (\phi \to \rho)))$ for an example. If we track the depth of every atom appearing on either side of the conditional, we find that the depth of all shared atoms in the antecedent are off by two from the corresponding atoms in the consequent. Switching formula order, parentheses order, and connectives doesn't resolve this issue.

Corollary 5. *Given an axiomatically appropriate constant specification* CS*, if* $\phi \in \mathbf{B}.\mathrm{J}_{\mathsf{CS}}$*, there is a justification term t such that for any depth substitution d and any* $n \in \mathbb{Z}$*,* $t{:}d(\phi, n) \in \mathbf{B}.\mathrm{J}_{\mathsf{CS}}$.

Proof. Let $\phi \in \mathbf{B}.\mathrm{J}_{\mathsf{CS}}$, d be a depth substitution, and $n \in \mathbb{Z}$. By Lemma 4, there is t so that $t{:}\ \phi \in \mathbf{B}.\mathrm{J}_{\mathsf{CS}}$. By Lemma 3, $d(t{:}\ \phi, n) \in \mathbf{B}.\mathrm{J}_{\mathsf{CS}}$. By the definition of depth substitution, $d(t{:}\phi, n) = t{:}d(\phi, n)$, and we are done. □

Note first the order of the quantifiers here: *there is* a t so that *for all* d and n... Note also that the choice of t did not rely on the choice of d or n. It's then an immediate observation that if t is any justification of a formula, then t is also a justification of any depth substitution of that formula at any depth. Thus, as a specific subcase, any available proof term for a theorem of **B** is also a proof term for every possible depth mapping of that theorem, which is again a theorem of **B**. And this may be expected, as depth substitutions (as with uniform substitutions) do not change the *structure* of their input formula. As we will see below, this result makes sense as it is a parallel notion to actual *proofs* of theorems in the base logic, which should be indistinguishable among theorems in which atoms have been simply swapped around (with respect to depth).

Finally, note that we could generalize to arbitrary logics **L**, as long as the following criteria are met.

(1) $\mathbf{L}.\mathrm{J}_{\mathsf{CS}}$ satisfies internalization,

(2) **L** is hyperformal, and

(3) Depth substitutions d are defined in such a way that they associate with proof terms: $d(t{:}\phi, n) = t{:}d(\phi, n)$.

Of the three available parameters, we will not budge on the definition of a depth substitution. Thus our result should apply to all hyperformal logics **L**, as long as we can guarantee condition (1). Fortunately, Standefer (2022) does much of the work for us in this regard, having proven that justification logics extending **B** up through **R**, with an AACS specified, maintain internalization. This leads us naturally to:

Lemma 6. *For subsystems* **L** *of* $\mathbf{DR}^-$ *at least as strong as* **B***, given an* AACS*, if* $\phi \in \mathbf{L}.\mathrm{J}_{\mathsf{CS}}$*, there is a justification term t so that for any depth substitution d and* $n \in \mathbb{Z}$*,* $t{:}d(\phi, n) \in \mathbf{L}.\mathrm{J}_{\mathsf{CS}}$.[7]

The proof for each logic thus reduces to proofs of internalization. Hyperformalism is easily taken care of, as strengthening the subsystem **L** within this range involves the removal or addition of inference rules and thus the removal or addition of justification axioms and rules to $\mathbf{B}.\mathrm{J}_{\mathsf{CS}}$, as noted in Section 2. The only additional case we did not cover in Theorem 4 is that of (JR6), but this rule is also easily seen to not violate depth relevance, as the premises and conclusion ultimately involve justifications of formulas which already depth share variables, being themselves members of

[7]We remind the reader here that the lemma is given for *subsystems*, and as given in Definition 3, these logics can be formulated only with a subset of the rules (R1)-(R5). It is clear that this restriction has some *ad hoc* flavor, but alas, we're doing proof theory in Hilbert systems. Some weirdness is bound to happen.

a base logic that enjoys drp. Thus nothing new is needed except to show that the additional base logic axioms are invariant under depth substitutions, which as before was shown in Logan (2022) up through $\mathbf{DR}^{-}$. The internalization proof is mildly different than Standefer's since our presentation of **B** eschews intensional connectives and the Ackermann constant, but it is nevertheless just as simple and no more groundbreaking.

6 First-Degree Proof Terms

The sentence $t{:}\phi$ may be read as one of many analogous natural language statements, but we will take "t records a proof of ϕ" as the preferred reading. The variables that make up t similarly record subproofs involved in the proof of ϕ. Naturally, one may wish to switch between results about proofs as objects over the base language **B** and results about justifications of theorems in **B**. We will call a justification term t a *first-degree proof term* (corresponding to a given constant specification CS) if, and only if, there is some $\phi \in \mathbf{B}$ for which $t{:}\phi \in \mathbf{B}.\mathsf{J}_{\mathsf{CS}}$. Here we present some notation for *what a proof is* and establish a correspondence between proofs and proof terms.

Definition 7. A *formula tree* over the language $\mathcal{L}^{+}$ is a directed graph that consists of edges attaching source nodes to destination nodes, subject to the following conditions.

- Source nodes can be $\mathcal{L}$-formulas, rule symbols, or justification constants
- Destination nodes can either be $\mathcal{L}$-formulas or rule symbols
- All source nodes have only one corresponding destination node
- If a destination node is a formula, it has *exactly* one corresponding source node
- If a destination node is a rule symbol, it has *at least* one corresponding source node
- All terminal nodes are formulas
- All initial nodes are either justification constants or subtrees

By TR, we refer to the set of all formula trees. A *subtree* corresponding to a formula tree Π is an abbreviated formula tree occurring within Π. A subtree may be empty, a justification-formula pair, a more complex formula tree, or the entire tree to which the subtree corresponds. For instance, the tree $c \to \phi$ has empty subtrees. An example of a formula tree is seen in Figure 1.

Formula trees glue together chains of formulas and rules, but an arbitrary formula tree does not necessarily record valid proofs in a given logic. The notation Λ_{ϕ} will be used to indicate that Λ is a formula tree with the single terminal node ϕ.

Definition 8. Let **L** be a subsystem of $\mathbf{DR}^{-}$ that contains **B**, and let CS be a constant specification. A *proof* corresponding to CS is a formula tree Π defined as follows.

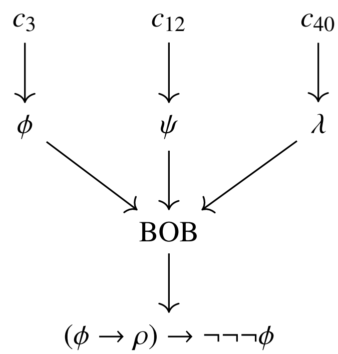
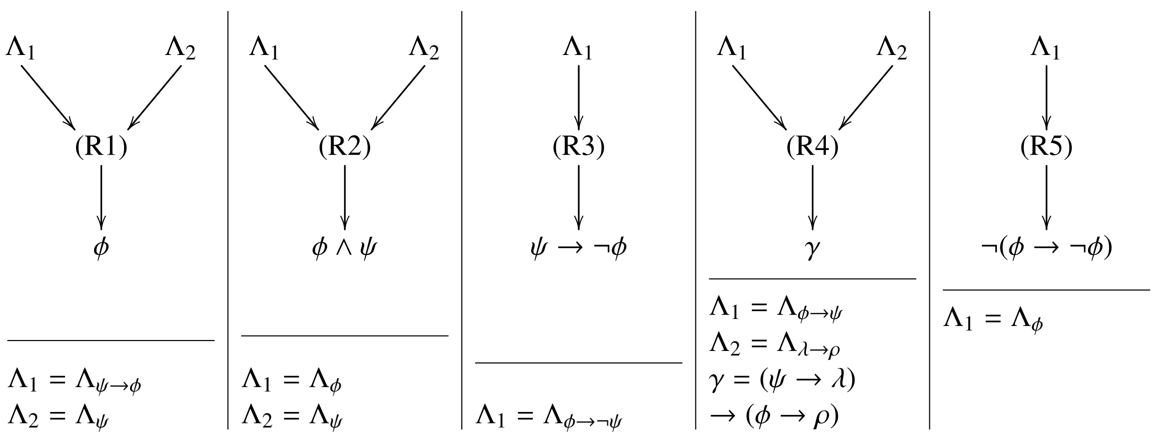


Figure 1: A formula tree with arbitrary rule node "BOB"

- If $\Pi = c \to \phi$, where ϕ is an instance of an axiom in $\mathbf{L}$ and $c{:}\phi \in \mathsf{CS}$, then the formula tree Π is a proof.
- If Λ_1 and Λ_2 are proofs, then the following are also proofs, provided the subproofs satisfy the given conditions.



If a proof Π terminates at ϕ and ϕ is a theorem of $\mathbf{L}$, we will say that Π is a *proof of* ϕ. By the *length* of the proof, we will mean the number of rule nodes that appear within it.

The syntax of a proof implies that proofs contain their subproofs. This behaves in roughly the same way that we have described proof terms. For a generic proof Π of a formula ϕ, we will interchange the notation $[\Lambda_1, \Lambda_2 \mid_R \phi]$, where Λ_1 and Λ_2 are (possibly empty) subtrees of Π and R is the (possibly empty) rule symbol connecting the subtrees to the formula.

Lemma 7. *Let* $\mathbf{L}$ *be a subsystem of* $\mathbf{DR}^-$ *that contains* $\mathbf{B}$ *and let* $\phi \in \mathbf{L}$. $\phi \in \mathbf{L}$ *if, and only if, there is a proof of* ϕ.

Proof. If $\phi \in \mathbf{L}$, then ϕ has a derivation. Proceed by induction on the complexity of this derivation. If ϕ is an axiom, then the formula tree ϕ is a proof. For the induction step, suppose now that all formulas in **L** with derivations of length $n-1$ have proofs, and that ϕ has derivation of length n. If the last derivation step was (R1), then there are some $\psi \to \phi \in \mathbf{L}$ and $\psi \in \mathbf{L}$. By induction hypothesis, these have respective proofs Π_1 and Π_2. Thus the formula tree $[\Pi_1, \Pi_2 \,|_{(R1)}\, \phi]$ is a proof of ϕ by Definition 8. The cases for (R2)-(R5) are proved similarly.

In the other direction, proceed by induction on the length of the proof Π of ϕ. If Π has length 0, it is the single-node formula tree ϕ and is an axiom of **L** by definition of proof. For the inductive step, assume the result holds for proofs of length $n-1$ and that the proof of ϕ is of length n. In one case, there are subproofs Π_1 of $\psi \to \phi$ and Π_2 of ψ so that Π is the proof $[\Pi_1, \Pi_2 \,|_{(R1)}\, \phi]$. By induction hypothesis, $\psi \to \phi \in \mathbf{L}$ and $\psi \in \mathbf{L}$. By (R1), $\phi \in \mathbf{L}$. Derivations are constructed similarly from the remaining forms of a proof. □

Lemma 7 establishes a correspondence between theorems of a logic and their proofs. It should be clear that this is not a bijective correspondence, as there can be many different proofs of any given theorem. One may, however, naturally conjecture that such a correspondence does connect proofs to proof terms, as opposed to formulas. It does seem plausible, at least, that rule nodes in formula trees encode a particular way of combining its subtrees, and that these encodings can be uniquely represented by the constants that appear in the tree's roots. We may then be able to translate between proofs and proof terms. We define a translation function $f : \mathsf{TR} \to \mathsf{JT}$ recursively.

- $f(c \to \phi) = c$
- If Λ_1, Λ_2 are proofs, then f is defined as follows
  - $f([\Lambda_1, \Lambda_2 \,|_{(R1)}\, \phi]) = [f(\Lambda_1) \cdot f(\Lambda_2)]$
  - $f([\Lambda_1, \Lambda_2 \,|_{(R2)}\, \phi]) = \mathfrak{a}(f(\Lambda_1), f(\Lambda_2))$
  - $f([\Lambda_1 \,|_{(R3)}\, \phi]) = \mathfrak{c}f(\Lambda_1)$
  - $f([\Lambda_1, \Lambda_2 \,|_{(R4)}\, \phi]) = \mathfrak{b}(f(\Lambda_1), f(\Lambda_2))$
  - $f(\Lambda_1 \,|_{(R5)}\, \phi) = *f(\Lambda_1)$
- Otherwise, for arbitrary formula trees of length m,

$$f([\Lambda_1, \ldots, \Lambda_m \,|_{\mathrm{R}}\, \phi]) = f(\Lambda_1) + \ldots + f(\Lambda_m)$$

The function f builds a proof term responsible for a given proof and forgets the formula it proves. We want to investigate how this function interacts with substitutions under composition, but first we need to extend our present definition of depth substitution.

Definition 9. A *proof substitution* corresponding to a depth substitution d is a function $d^p : \mathsf{TR} \times \mathbb{Z} \to \mathsf{TR}$ defined as follows. Let $\Pi = [\Lambda_1, \Lambda_2, \ldots, \Lambda_m \,|_{\mathrm{R}}\, \phi]$, **L** be a subsystem of $\mathbf{DR}^-$ with $\mathbf{B} \subseteq \mathbf{L}$, and $n \in \mathbb{Z}$.

- If $\Pi = c \to \phi$, then $d^p(\Pi, n) = c \to d(\phi, n)$

- If R represents a rule in **L** and Π is a proof of ϕ, then $1 \leq m \leq 2$ and R corresponds with one of the following.
  - (R1): $d^p(\Pi, n) = [d^p(\Lambda_1, n-1), d^p(\Lambda_2, n) \mid_{(R1)} d(\phi, n)]$, where Λ_1 is a proof of $\psi \rightarrow \phi$ and Λ_2 is a proof of ψ for some ψ
  - (R2): $d^p(\Pi, n) = [d^p(\Lambda_1, n), d^p(\Lambda_2, n) \mid_{(R2)} d(\phi, n)]$, where Λ_1 is a proof of λ_1, Λ_2 is a proof of λ_2, and $\phi = \lambda_1 \wedge \lambda_2$
  - (R3): $d^p(\Pi, n) = [d^p(\Lambda_1, n) \mid_{(R3)} d(\phi, n)]$, where Λ_1 is a proof of $\lambda_1 \rightarrow \neg\lambda_2$ and $\phi = \lambda_2 \rightarrow \neg\lambda_1$
  - (R4): $d^p(\Pi, n) = [d^p(\Lambda_1, n+1), d^p(\Lambda_2, n+1) \mid_{(R4)} d(\phi, n)]$, where Λ_1 is a proof of $\lambda_{11} \rightarrow \lambda_{12}$, Λ_2 is a proof of $\lambda_{21} \rightarrow \lambda_{22}$, and $\phi = (\lambda_{12} \rightarrow \lambda_{21}) \rightarrow (\lambda_{11} \rightarrow \lambda_{22})$
  - (R5): $d^p(\Pi, n) = [d^p(\Lambda_1, n+1) \mid_{(R5)} d(\phi, n)]$, where Λ_1 is a proof of λ and $\phi = \neg(\lambda \rightarrow \neg\lambda)$
- Otherwise, substitute subtrees in place:

$$d^p(\Pi, n) = [d^p(\Lambda_1, n), d^p(\Lambda_2, n), \ldots, d^p(\Lambda_m, n) \mid_R d(\phi, n)]$$

Lemma 8. *Given a depth substitution d, proof substitutions corresponding to d are unique.*

Proof. By induction on the length of a proof to which two proof substitutions corresponding to d are applied.

□

Lemma 9. *Let* **L** *be a subsystem of* $\mathbf{DR}^-$ *such that* $\mathbf{B} \subseteq \mathbf{L}$ *and* d^p *be a proof substitution. If* Π *is a proof of a theorem* $\phi \in \mathbf{L}$*, then for any* $n \in \mathbb{Z}$*,* $d^p(\Pi, n)$ *is a proof of* $d(\phi, n) \in \mathbf{L}$*.*

Proof. Proceed by induction on the length of the proof. If Π is a proof of an axiom ϕ, then $d^p(\Pi, n) = d(\phi, n)$. By Theorem 2 and our choice of logic, $d(\phi, n) \in \mathbf{L}$. Moreover, depth substitutions of axioms are themselves axioms, and an axiom taken as a formula tree is a proof of itself. Thus $d^p(\Pi, n)$ is a proof.

Suppose now that Π is a proof containing m rule nodes and the theorem is true of proofs containing fewer than m. We give two cases and leave the rest to the reader. If R = (R2), then for some formula λ_1 and λ_2, $\phi = \lambda_1 \wedge \lambda_2$ and Π is of the form $[\Lambda_{\lambda_1}, \Lambda_{\lambda_2} \mid_R \lambda_1 \wedge \lambda_2]$, where Λ_{λ_1} and Λ_{λ_2} are proofs. By the induction hypothesis, $d^p(\Lambda_{\lambda_1}, n)$ and $d^p(\Lambda_{\lambda_2}, n)$ are proofs of, say, $d(\lambda_1, n)$ and $d(\lambda_2, n)$ respectively. Thus

$$\begin{aligned} d^p(\Pi, n) &= [d^p(\Lambda_{\lambda_1}, n), d^p(\Lambda_{\lambda_2}, n) \mid_{(R2)} d(\lambda_1 \wedge \lambda_2, n)]] \\ &= [[\ldots \mid_S d(\lambda_1, n)], [\ldots \mid_T d(\lambda_2, n)] \mid_{(R2)} d(\lambda_1 \wedge \lambda_2, n)]] \\ &= [[\ldots \mid_S d(\lambda_1, n)], [\ldots \mid_T d(\lambda_2, n)] \mid_{(R2)} d(\lambda_1, n) \wedge d(\lambda_2, n)]] \end{aligned}$$

is a proof of $d(\phi, n)$, where the ellipses indicate nested subtrees of the subtrees within which they appear and S and T are other rule symbols.

If instead R = (R1), then $\Pi = [\Lambda_{\psi\to\phi}, \Lambda_\psi \mid_{(R1)} \phi]$ where $\Lambda_{\psi\to\phi}$ and Λ_ψ are proofs. By induction hypothesis, $d^p(\Lambda_{\psi\to\phi}, n-1)$ is a proof of $d(\psi \to \phi, n-1)$ and $d^p(\Lambda_\psi, n)$ is a proof of $d(\psi, n)$.

$$\begin{aligned} d^p(\Pi, n) &= [d^p(\Lambda_1, n-1), d^p(\Lambda_2, n) \mid_{(R1)} d(\phi, n)] \\ &= [[\ldots \mid_S d(\psi \to \phi, n-1)], [\ldots \mid_T d(\psi, n)] \mid_{(R1)} d(\phi, n)]] \\ &= [[\ldots \mid_S d(\psi, n) \to d(\phi, n)], [\ldots \mid_T d(\psi, n)] \mid_{(R1)} d(\phi, n)]] \end{aligned}$$

Hence $d^p(\Pi, n)$ is a proof of $d(\phi, n)$.

□

The statement of Lemma 8 indicates that the choice of depth substitution completely determines the behavior of the corresponding proof substitution, i.e. that this function is well-defined. But, moreover, Lemma 9 implies that, given a fixed theorem, our initial choice of proof doesn't really matter—we end up with a proof of the same depth substitutiton instance, regardless of which proof we selected. However, what this lemma fails to address is the structure of the proofs we start and end with. How can we ascertain that these two proofs don't drastically differ from each other?

We'll show just this. Recall that our function f isolates first-degree proof terms in a justification formula corresponding to a proof of that formula.

Lemma 10. *Let d be a depth substitution and d^p be the unique proof substitution associated with d. Then $f \circ d^p$ and f are pointwise identical on formula trees for any $n \in \mathbb{Z}$.*

Proof. Proceed by induction on the length of the formula tree Π. If $\Pi = c \to \phi$, then $f(d^p(c \to \phi, n)) = f(c \to d(\phi, n)) = c = f(c \to \phi)$. Assume now that the lemma holds for formula trees of length $k-1$ and that Π has length k.

If $\Pi = [\Lambda_1, \ldots, \Lambda_m \mid_R \phi]$ is not a proof, then

$$\begin{aligned} f(d^p(\Pi, n)) &= f([d^p(\Lambda_1, n), \ldots, d^p(\Lambda_m, n) \mid_R d(\phi, n)]) \\ &= f(d^p(\Lambda_1, n)) + \ldots + f(d^p(\Lambda_m, n)) \\ &= f(\Lambda_1) + \ldots + f(\Lambda_m) && \text{(by IH)} \\ &= f(\Pi) \end{aligned}$$

If Π is a proof, however, then several cases follow. We highlight just one case and leave the rest to the reader. If $\Pi = [\Lambda_1, \Lambda_2 \mid_{(R1)} \phi]$,

$$\begin{aligned} f(d^p([\Lambda_1, \Lambda_2 \mid_{(R1)} \phi], n)) &= f([d^p(\Lambda_1, n-1), d^p(\Lambda_2, n) \mid_{(R1)} d(\phi, n)]) \\ &= [f(d^p(\Lambda_1, n-1)) \cdot f(d^p(\Lambda_2, n))] \\ &= [f(\Lambda_1) \cdot f(\Lambda_2)] && \text{(by IH)} \\ &= f(\Pi) \end{aligned}$$

□

What we have just seen then is that, indeed, first-degree proof terms are invariant under proof substitutions, and this generally reveals that, while formulas of the base logic are taken to depth substitution instances of themselves, the structure of the proofs underlying the original formula and its substitution does not change one bit.

7 Final Remarks

In this paper, we showed that we can make a minimal, reasonable set of assumptions about proof terms over a relevant logic and effortlessly recover the hyperformal characteristic of the original logic. The crux of this examination is found in Corollary 5, in which we determined for each theorem of the justification logic, there is at least one proof term which justifies every depth substitution instance of it. We then examined the correspondence between first-degree justifications and proofs of theorems, which allowed us to draw reasonable corollaries about the nature of proofs in logics weaker than $\mathbf{DR}^{-}$ and that any depth substitution of a theorem can be proved through the same exact proof structure as the proof of the original theorem, no matter how arbitrarily atoms may be mapped. Thus whenever we commit to an axiomatically appropriate constant specification, we obtain a rich modal structure over hyperformal logics in which prior results about hyperformalism are mirrored in the structures of proofs.

The invariance of theoremhood in **B** (and other logics considered in this paper) under depth substitution can be seen as a *feature* of the theorems in **B** and the strength of the relevant conditional. Formulas on either side of the conditional do not show up by surprise, as it were, but rather the positions in which they appear have some predictable relation to their conditional counterpart. That proof substitutions are also invariant can be seen to communicate the strength and restriction of inference in **B**, and it indicates that depth substitutions do not take theorems with a given connective structure "very far" from that structure. This structural robustness is intriguing, but it may alas be a symptom of a stronger phenomenon taking place in the limitations placed on inference in the logic.

Finally, we make a small note on a recent class of substitutions introduced in a forthcoming article from Ferguson and Logan. These provide a stronger relevance property—dubbed there as *cn-relevance*—than that of depth relevance and are defended on the grounds of topic transparency. The analogous substitutions act on atomic formulas and sequences of operators, as opposed to integers representing depth. It is found there that the logics **B** and **BM** enjoy *cn-relevance*, in a similar sense to our definition for the enjoyment of depth relevance, yet marginally stronger logics such as **DW** provide immediate counterexamples that do not enjoy this novel relevance criterion. It is quite simple to take the scaffolding of this paper and adopt it instead to this other class of substitutions, of which depth is just a subclass. The main idea is then that reduced sequence substitutions and justification terms do not interact, passing through each other just as transparently as depth substitutions and justification terms did. This is again an epistemic choice which we find justified in light of a proper choice of constant specification. The exploration of the interactions between these substitution classes and justifications, however, remains outside the scope of this paper, and the present author would encourage such exploration if the reader should be interested.

Acknowledgements

This topic was introduced to me by Shay Allen Logan, and I am indebted to him for helping me through many drafts of this project. A first draft was presented to the Perspectives on Logic and Philosophy conference at Ruhr University Bochum in 2023, and I owe Andrew Tedder and Heinrich Wansing thanks for their helpful comments during and after my talk, as they steered my results in fruitful directions. I additionally would like to thank Shawn Standefer for pre-reviewing a revised draft of the paper and providing helpful and detailed feedback. Finally, the K-State Logic Group at Kansas State University sat through multiple presentations of this material, and it is due to their patient and thorough comments that this paper was submitted and eventually accepted into this journal.

References

Anderson, A. R. and N. D. Belnap (1975). *Entailment: The Logic of Relevance and Neccessity, Vol. I*. Princeton University Press.

Artemov, S. and M. Fitting (2019). *The Basics of Justification Logic*, pp. xii, 11–30. Cambridge Tracts in Mathematics. Cambridge University Press.

Belnap, N. D. (1960). Entailment and relevance. *The Journal of Symbolic Logic 25*(2), 144–146. https://doi.org/10.2307/2964210.

Brady, R. T. (1984). Depth relevance of some paraconsistent logics. *Studia Logica: An International Journal for Symbolic Logic 43*(1/2), 63–73. https://doi.org/10.1007/BF00935740.

Ferguson, T. M. and S. A. Logan (forthcoming). Topic transparency and variable sharing in weak relevant logics. *Erkenntnis*, 1–28.

Fitting, M. (2017). Paraconsistent logic, evidence, and justification. *Studia Logica 105*(6), 1149–1166. https://doi.org/10.1007/s11225-017-9714-3.

Logan, S. A. (2021). Strong depth relevance. *Australasian Journal of Logic 18*(6), 645–656. https://doi.org/10.26686/ajl.v18i6.7081.

Logan, S. A. (2022). Depth relevance and hyperformalism. *Journal of Philosophical Logic 51*(4), 721–737. https://doi.org/10.1007/s10992-021-09648-y.

Mares, E. D. (2004). *Disjunctive syllogism*, pp. 176–188. Cambridge University Press. https://doi.org/10.1017/CBO9780511520006.011.

Méndez, J. M. and G. Robles (2012). A general characterization of the variable-sharing property by means of logical matrices. *Notre Dame Journal of Formal Logic 53*(2), 223 – 244. https://doi.org/10.1215/00294527-1715707.

Robles, G. and J. M. Méndez (2014). Generalizing the depth relevance condition: Deep relevant logics not included in R-mingle. *Notre Dame Journal of Formal Logic 55*(1), 107–127. https://doi.org/10.1215/00294527-1960461.

Standefer, S. (2022). Weak relevant justification logics. *Journal of Logic and Computation*. https://doi.org/10.1093/logcom/exac057.

Szmuc, D. E. (2021). The (greatest) fragment of classical logic that respects the variable-sharing principle (in the fmla-fmla framework). *Bulletin of the Section of Logic 50*(4), 421–453. https://doi.org/10.18778/0138-0680.2021.08.