– Accepted for publication in Logique & Analyse –

A non-classical refinement of the interpolation property for classical propositional logic

Peter Milne

Abstract

We refine the interpolation property of the $\{\wedge, \vee, \neg\}$-fragment of classical propositional logic, showing that if $\nvDash \neg\phi$, $\nvDash \psi$ and $\phi \vDash \psi$ then there is an interpolant χ, constructed using at most atomic formulas occurring in both ϕ and ψ and negation, conjunction and disjunction, such that (i) ϕ entails χ in Kleene's strong three-valued logic and (ii) χ entails ψ in Priest's Logic of Paradox.

Keywords: Interpolation theorem for classical propositional logic · Kleene's strong 3-valued logic · Priest's Logic of Paradox

1 Introduction

Suppose that ϕ classically entails ψ, that ϕ is not a classical contradiction and that ψ is not a classical tautology. Then ϕ and ψ must share non-logical vocabulary, for else one could make ϕ true and ψ false at the same time. That there must be *some* overlap in non-logical vocabulary between premise and conclusion is obvious. Possession of the interpolation property takes this line of thought much further:

> If $\nvDash \neg\phi$, $\nvDash \psi$ and $\phi \vDash \psi$ then there is a formula χ containing only atomic formulas common to ϕ and ψ and such that $\phi \vDash \chi$ and $\chi \vDash \psi$.[1]

[1]Surprisingly few textbooks prove this theorem. (Hodges 2001) and (Hunter 1971) are exceptions. They prove it in the slightly stronger form

> if $\phi \vDash \psi$ and at least one atomic formula is common to ϕ and ψ then there is a formula χ containing only atomic formulas common to ϕ and ψ and such that $\phi \vDash \chi$ and $\chi \vDash \psi$.

Now, as is well known, many logics fail to possess the interpolation property—see, *e.g.*, (Schumm 1986) and the references therein, (Renardel de Lavalette 1981), (Urquhart 1993), (Roorda 1994). On the other hand, many logics do share the interpolation properties exhibited by classical propositional logic or close analogues thereof. Two that do are Kleene's strong three-valued logic ($K3$; Kleene 1952, §64) and Priest's Logic of Paradox (LP; Priest 1979). To be exact, abbreviating 'atomic formula' to 'atom' we have

> if $\phi \nvDash_{K3}$ and $\phi \vDash_{K3} \psi$ then there is a formula χ containing only atoms common to ϕ and ψ and such that $\phi \vDash_{K3} \chi$ and $\chi \vDash_{K3} \psi$[2]

and

> if $\nvDash_{LP} \psi$ and $\phi \vDash_{LP} \psi$ then there is a formula χ containing only atoms common to ϕ and ψ and such that $\phi \vDash_{LP} \chi$ and $\chi \vDash_{LP} \psi$.

That Kleene's logic has this interpolation property was shown by Kamila Bendová (2005).[3] That Priest's logic has the stated interpolation property follows immediately *via* the *Duality Principle* that links $K3$ and LP, namely,

$$\neg\phi \vDash_{K3} \neg\psi \text{ iff } \psi \vDash_{LP} \phi \textit{ and } \neg\phi \vDash_{K3} \text{ iff } \vDash_{LP} \phi,$$

and the fact that Double Negation Introduction and Elimination are sound in $K3$.[4]

Although $K3$ and LP have the indicated interpolation properties which are the minimal modifications of the classical case to, respectively, a logic without

[2]For a logic X, by $\phi \vDash_X$ we mean that $\phi \vDash_X \xi$ for *all* ξ and call ϕ an *anti-theorem* of the logic X. Classically, $\phi \vDash$ iff $\vDash \neg\phi$; in $K3$, $\phi \vDash_{K3}$ iff $\phi \vDash_{K3} \neg\phi$. LP has no anti-theorems.

[3]A proof that propositional $K3$ possesses the interpolation property, different in approach to Bendová's in that it is modelled on the proof for classical propositional logic in (Hunter 1971), is to be found in the Appendix.

[4]The interpolation theorem for classical propositional logic is equivalent to the following: if $\neg\phi$ and ψ are not classical tautologies but $\neg\phi \vee \psi$ is then there is a formula χ containing only non-logical vocabulary common to ϕ and ψ such that $\neg\phi \vee \chi$ and $\neg\chi \vee \psi$ are both classical tautologies. Since all and only classical tautologies are theorems of the Logic of Paradox, we immediately obtain one form of interpolation theorem for LP. But this does *not* translate back into the form given in the text because, as Priest notes (Priest 1979, §III.9), LP lacks Disjunctive Syllogism.

Beall *et al.* (2013, p. 18, n. 5) say, 'As the referee noted, Takano's result in (Takano 1989) immediately delivers interpolation for LP.' If so, it is in the form just noted, *i.e.*, for *theorems* of LP of some fixed logical form. Takano provides a condition that is sufficient on the assumption of expressive completeness, namely, in the case of LP, that there be some function $f : \{0, 1/2, 1\}^2 \to \{0, 1/2, 1\}$ which satisfies the constraints (i) that, for any $x, y, z \in \{0, 1/2, 1\}$, if $f(x, y) \geq 1/2$ and $f(y, z) \geq 1/2$ then $f(x, z) \geq 1/2$ and (ii) that, for all $x, y \in \{0, 1/2, 1\}$, $f(x, y) \geq 1/2$ or $f(y, x) \geq 1/2$. As is clear from the truth-tables in Section 2 in the light of Lemma 1, any formula built up from atoms using at most $\wedge$, $\vee$, and $\neg$ that expresses a function satisfying these constraints is a classical tautology in two propositional variables and, classical tautologies being theorems of the Logic of Paradox, the resulting "interpolation property" is entirely trivial.

theorems ($K3$) and a logic without anti-theorems (LP), it's worth noting that the logic whose valid inferences are exactly those common to $K3$ and LP does not. This logic, called Kalman implication in (Makinson 1973), has neither theorems (since $K3$ has none) nor anti-theorems (since LP has none), but since $\phi \wedge \neg\phi$ is (classically hence) $K3$-unsatisfiable and $\psi \vee \neg\psi$ is a (classical hence) LP logical truth, $\phi \wedge \neg\phi \vDash_{Kalman} \psi \vee \neg\psi$ for any ϕ and ψ.[5]

In the classical case, as ϕ is not a contradiction, any interpolant χ is likewise not a contradiction and, since ψ is not a tautology, χ is not a tautology. That aside, the statement of the interpolation theorem gives us little to go on. Here I prove a refinement of the theorem for the $\{\wedge, \vee, \neg\}$-fragment of classical propositional logic which tells us a little more: we can construct an interpolant that is entailed by ϕ in $K3$ and entails ψ in LP. This is not a trivial consequence of the interpolation properties of Kleene's and Priest's propositional logics, not least because there are classically valid entailments that hold in neither $K3$ nor LP. $\neg\phi \wedge (\phi \vee \psi) \vDash \neg\chi \vee (\chi \wedge \psi)$ is one such. The result is more than a mere novelty, for, arguably, $K3$ is what survives of classical propositional logic when one gives up the joint exhaustiveness of truth and falsity and LP is what survives when one gives up their mutual exclusivity.

2 Kleene's and Priest's three-valued logics

Kleene's strong three-valued logic and Priest's Logic of Paradox share the same three-valued truth-tables for negation, conjunction and disjunction. These are:

ϕ	$\neg\phi$
1	0
1/2	1/2
0	1

$\phi \wedge \psi$	1	1/2	0
1	1	1/2	0
1/2	1/2	1/2	0
0	0	0	0

$\phi \vee \psi$	1	1/2	0
1	1	1	1
1/2	1	1/2	1/2
0	1	1/2	0

.

The logics differ in that Kleene's takes 1 to be the only designated value, Priest's takes 1 and 1/2 to be designated.

Following the notation of (Priest 1979), an assignment u of values to atomic formulas extends to a valuation u^+ on all formulas constructed employing negation, conjunction and disjunction in accordance with the truth-tables. We consider no other formulas in what follows and no other valuations. ($^+$ is an injective function from assignments to valuations.) We restrict attention to inferences with at most a single premise and take '$\vDash$' to stand for classical consequence, '$\vDash_{K3}$' to stand for preservation of the value 1 from premise to conclusion under all valuations, and '$\vDash_{LP}$' to stand for preservation of the value 0 from conclusion to premise under all valuations.[6]

[5]Bendová (2005, p. 130, Remark 1) notes that requiring premise and conclusion to share an atom does not remove all failures of interpolation in Kalman implication.

[6]$\phi \vDash_{Kalman} \psi$ iff, under all valuations u^+, $u^+(\phi) \leq u^+(\psi)$.

The Duality Principle stated above is an immediate consequence of these definitions.

Looking at the truth-tables two points leap to the eye and can be shown more formally by induction on length of formula:

- an assignment which assigns only 0 and/or 1 to atoms extends to a valuation which assigns only 0 and 1 fully in classical fashion;
- an assignment which assigns only $1/2$ to atoms extends to a valuation which assigns only $1/2$.

This second observation shows that Kleene's logic ($K3$) has no theorems—no formula must take the value 1—and that Priest's logic (LP) has no anti-theorems—no formula must take the value 0.

A third fact is almost equally obvious:

Lemma 1 Given assignments u and v, if v differs from u only in assigning the value $1/2$ to one or more atoms to which u assigns either 0 or 1, then, for all formulas ϕ, $v^+(\phi) = u^+(\phi)$ or $v^+(\phi) = 1/2$.

Proof The claim holds trivially for atoms. Suppose that it holds for all formulas with k or fewer occurrences of connectives, and let ϕ contain $k+1$ occurrences of connectives. Proof is by cases.

- ϕ is $\neg\psi$. Then $v^+(\phi) = 1 - v^+(\psi)$ and, by the induction hypothesis, $v^+(\psi) = u^+(\psi)$ or $v^+(\psi) = 1/2$. Consequently, $v^+(\phi) = 1 - u^+(\psi) = u^+(\phi)$ or $v^+(\phi) = 1 - 1/2 = 1/2$.
- ϕ is $\psi \wedge \chi$. Then $v^+(\phi) = \min\{v^+(\psi), v^+(\chi)\}$ and, by the induction hypothesis, $v^+(\psi) = u^+(\psi)$ or $v^+(\psi) = 1/2$ and $v^+(\chi) = u^+(\chi)$ or $v^+(\chi) = 1/2$. Consequently, $v^+(\phi) = 1 = u^+(\phi)$ if $v^+(\psi) = u^+(\psi) = 1$ and $v^+(\chi) = u^+(\chi) = 1$, $v^+(\phi) = 0 = u^+(\phi)$ if $v^+(\psi) = u^+(\psi) = 0$ or $v^+(\chi) = u^+(\chi) = 0$, and otherwise $v^+(\phi) = 1/2$.
- ϕ is $\psi \vee \chi$. Then $v^+(\phi) = \max\{v^+(\psi), v^+(\chi)\}$ and, by the induction hypothesis, $v^+(\psi) = u^+(\psi)$ or $v^+(\psi) = 1/2$ and $v^+(\chi) = u^+(\chi)$ or $v^+(\chi) = 1/2$. Consequently, $v^+(\phi) = 0 = u^+(\phi)$ if $v^+(\psi) = u^+(\psi) = 0$ and $v^+(\chi) = u^+(\chi) = 0$, $v^+(\phi) = 1 = u^+(\phi)$ if $v^+(\psi) = u^+(\psi) = 1$ or $v^+(\chi) = u^+(\chi) = 1$, and otherwise $v^+(\phi) = 1/2$. □

An immediate consequence of this last lemma will do some work in what follows:

Corollary 1.1 If a valuation u^+ assigns only the values 0 and/or 1 to the members of some set of formulas X then any valuation which agrees with u^+ on those atoms occurring in members of X to which the latter assigns 0 and 1 agrees with u^+ on all members of X. □

This corollary yields a criterion for the existence of classical (two-valued) valuations on sets of formulas:

Observation 1 If a valuation u^+ assigns only the values 0 and/or 1 to the members of some set of formulas X, there is a *classical* valuation which agrees with u^+ on those atoms occurring in members of X to which the latter assigns 0 and 1 and agrees with u^+ on all members of X.

From this it follows that all classical tautologies are theorems of LP (*cf.* Priest 1979, §III.8) and all classical contradictions are anti-theorems of $K3$.

3 The refinement

Lemma 2 If $\phi \vDash \psi$ then, when the valuation u^+ assigns the value 1 to ϕ and the valuation v^+ assigns the value 0 to ψ, there is an atom p, common to ϕ and ψ such that either $u(p) = 1$ and $v(p) = 0$ or $u(p) = 0$ and $v(p) = 1$.

Proof $u \neq v$, for were the same valuation to assign 1 to ϕ and 0 to ψ, by Observation 1 there would be a purely classical valuation making ϕ true and ψ false, contrary to hypothesis.

u and v must, then, disagree but were their disagreement limited to these four forms—

- u and v disagree on one or more atoms that occur in ϕ but not in ψ
- u and v disagree on one or more atoms that occur in ψ but not in ϕ
- u assigns the value $1/2$ to one or more atoms common to ϕ and ψ to which v assigns either 0 or 1
- v assigns the value $1/2$ to one or more atoms common to ϕ and ψ to which u assigns either 0 or 1

—we could define this assignment w:

- w agrees with u on atoms that occur in ϕ but not in ψ;
- w agrees with v on atoms that occur in ψ but not in ϕ;

- w assigns what v assigns when u assigns the value $1/2$ to one or more atoms common to ϕ and ψ to which v assigns either 0 or 1;
- w assigns what u assigns when v assigns the value $1/2$ to one or more atoms common to ϕ and ψ to which u assigns either 0 or 1;
- elsewhere w agrees with u and v.

w extends to the valuation w^+. As w differs from u on atoms occurring in ϕ at most by assigning 0 or 1 where u assigns $1/2$, by Corollary 1.1, $w^+(\phi) = u^+(\phi) = 1$. Similarly, as w differs from v on atoms occurring in ψ at most by assigning 0 or 1 where v assigns $1/2$, by Corollary 1.1 again, $w^+(\psi) = v^+(\psi) = 0$. And then, by Observation 1 and contrary to hypothesis, there is a classical valuation under which ϕ is true and ψ is false.

Consequently, as claimed, there must be at least one atom p, common to ϕ and ψ such that either $u(p) = 1$ and $v(p) = 0$ or $u(p) = 0$ and $v(p) = 1$. □

Theorem 1 If $\nvDash \neg\phi$, $\nvDash \psi$ and $\phi \vDash \psi$ then there is a formula χ containing only atoms common to ϕ and ψ and such that $\phi \vDash_{K3} \chi$ and $\chi \vDash_{LP} \psi$.

Proof Firstly, since $\nvDash \neg\phi$ and $\nvDash \psi$, ϕ and ψ have at least one atom in common. Secondly, as $\nvDash \neg\phi$, there is a classical valuation u^+ such that $u^+(\phi) = 1$; likewise, as $\nvDash \psi$, there is a classical valuation v^+ such that $v^+(\psi) = 0$. In the light of this and Lemma 2, any valuation assigning either 1 to ϕ or 0 to ψ must assign either 0 or 1 to an atom common to ϕ and ψ.

Let $p_1, p_2, \ldots, p_n$ be the atoms common to ϕ and ψ. Corresponding to an assignment u, let χ_u be a conjunction of literals containing p_i if $u(p_i) = 1$, $\neg p_i$ if $u(p_i) = 0$, and simply ignoring p_i if $u(p_i) = 1/2$. By what has just been observed, there is always at least one conjunct when a valuation assigns either 1 to ϕ or 0 to ψ.

For each assignment u of values to just the atoms occurring in one or other or both of ϕ and ψ, label it with a 1 if $u^+(\phi) = 1$, label it with a 0 if $u^+(\psi) = 0$. (As $\phi \vDash \psi$, each assignment bears at most one label.) Let χ_1 be the disjunction of those conjunctions χ_u for which u is labelled with a 1 and let χ_0 be the disjunction of those conjunctions χ_u for which u is labelled with a 0.

Now, let u^+ be a valuation for which $u^+(\phi) = 1$ and let u be the induced assignment to just the atoms occurring in one or other or both of ϕ and ψ. By construction, $u^+(\chi_u) = 1$ and hence $u^+(\chi_1) = 1$. Thus $\phi \vDash_{K3} \chi_1$. Next, let v^+ be a valuation for which $v^+(\psi) = 0$ and let v be the induced assignment to just the atoms occurring in one or other or both of ϕ and ψ. It remains to show that $v^+(\chi_1) = 0$. By Lemma 2, for any valuation w^+, inducing the assignment w on the atoms occurring in one or other or both of ϕ and ψ, for which $w^+(\phi) = 1$ there is an atom p, common to ϕ and ψ, such that either $w(p) = 1$ and $v(p) = 0$ or $w(p) = 0$ and $v(p) = 1$. In the first case, χ_w contains p and $v^+(\chi_w) = 0$; in the second, χ_w contains $\neg p$ and again $v^+(\chi_w) = 0$. As this holds for all assignments w labelled with a 1, $v^+(\chi_1) = 0$. Thus $\chi_1 \vDash_{LP} \psi$.

We have shown that

$$\phi \vDash_{K3} \chi_1 \text{ and } \chi_1 \vDash_{LP} \psi.$$

$\neg\chi_0$ (and thus $\chi_1 \wedge \neg\chi_0$ and $\chi_1 \vee \neg\chi_0$) serves equally well as an interpolant with the properties we seek. □

4 Conclusion

We have strengthened one form of the interpolation theorem for the $\{\wedge, \vee, \neg\}$-fragment of classical propositional logic: when $\phi \vDash \psi$ and when ϕ is not a classical contradiction and ψ is not a classical tautology, we have shown how to construct an interpolant χ such that

$$\phi \vDash_{K3} \chi \text{ and } \chi \vDash_{LP} \psi.$$

This cannot be further strengthened so as to read

> if $\phi \vDash \psi$ and at least one atom is common to ϕ and ψ then there is a formula χ containing only atoms common to ϕ and ψ and such that $\phi \vDash_{K3} \chi$ and $\chi \vDash_{LP} \psi$.

$(\phi \wedge \neg\phi) \vee \psi \vDash (\phi \vee \neg\phi) \vee \chi$ but when ϕ, ψ and χ are distinct atoms, no formula containing ϕ as sole atom is a $K3$-consequence of $(\phi \wedge \neg\phi) \vee \psi$.

Acknowledgement

I thank a referee for advice on the presentation of this material and for drawing (Bendová 2005) to my attention.

Appendix: Interpolation Theorem for $K3$

Theorem 2 If $\phi \nvDash_{K3}$ and $\phi \vDash_{K3} \psi$ then there is a formula χ employing only atoms common to ϕ and ψ such that $\phi \vDash_{K3} \chi$ and $\chi \vDash_{K3} \psi$.

Proof is by induction on the number of atoms occurring in ϕ that do not occur in ψ. If that number is zero, ϕ itself serves as an interpolant, and this provides the base case of the induction. Now suppose that an interpolant exists when at most m atoms occur in the premise but not in the conclusion—the induction hypothesis—and suppose too that no more than $m+1$ atoms occur in ϕ but not in ψ. Let q be one of these and let $p_1, p_2, \ldots, p_n$ be all other atoms occurring in ϕ, at most m of which do not occur in ψ.

Let ϕ_1 be the formula obtained by replacing all occurrences of q in ϕ with $(p_1 \vee \neg p_1) \vee (p_2 \vee \neg p_2) \vee \ldots \vee (p_n \vee \neg p_n)$; let ϕ_2 be the formula obtained by replacing all occurrences of q in ϕ with $(p_1 \wedge \neg p_1) \wedge (p_2 \wedge \neg p_2) \wedge \ldots \wedge (p_n \wedge \neg p_n)$. We show, first, that $\phi_1 \vee \phi_2 \vDash_{K3} \psi$, then that $\phi \vDash_{K3} \phi_1 \vee \phi_2$.

A valuation only assigns 1 to $\phi_1 \vee \phi_2$ if it assigns 1 to one or other of ϕ_1 and ϕ_2. Suppose first, then, that it assigns 1 to ϕ_1. It can only do this by assigning 0 or 1 to at least one of $p_1, p_2, \ldots, p_n$, thus assigning 1 to $(p_1 \vee \neg p_1) \vee (p_2 \vee \neg p_2) \vee \ldots \vee (p_n \vee \neg p_n)$; the valuation that differs at most by assigning 1 to q also assigns 1 to ϕ, and hence to ψ; as q doesn't occur in ψ, the original valuation must also assign 1 to ψ. Suppose next that our original valuation assigns 1 to ϕ_2. It can only do this by assigning 0 or 1 to at least one of $p_1, p_2, \ldots, p_n$, thus assigning 0 to $(p_1 \wedge \neg p_1) \wedge (p_2 \wedge \neg p_2) \wedge \ldots \wedge (p_n \wedge \neg p_n)$; the valuation that differs at most by assigning 0 to q also assigns 1 to ϕ, and hence to ψ; again, as q doesn't occur in ψ, the original valuation must also assign 1 to ψ. Hence $\phi_1 \vee \phi_2 \vDash_{K3} \psi$.

It remains to show that $\phi \vDash_{K3} \phi_1 \vee \phi_2$. Suppose that an assignment to p_1, p_2, …, p_n and q leads to assigning 1 to ϕ. At least one of the atoms must be assigned 0 or 1. Were q alone to take one of these values, as it does not occur in ψ the assignment which assigns $1/2$ to all atoms save q leads to ψ itself taking the value $1/2$, contradicting our starting point that $\phi \vDash_{K3} \psi$. So at least one among $p_1, p_2, \ldots, p_n$ takes one of the values 0 and 1 and hence $(p_1 \vee \neg p_1) \vee (p_2 \vee \neg p_2) \vee \ldots \vee (p_n \vee \neg p_n)$ takes the value 1 and $(p_1 \wedge \neg p_1) \wedge (p_2 \wedge \neg p_2) \wedge \ldots \wedge (p_n \wedge \neg p_n)$ takes the value 0. If 1 is assigned to q, ϕ_1 and hence $\phi_1 \vee \phi_2$ takes the value 1; if 0 is assigned to q, ϕ_2 and hence $\phi_1 \vee \phi_2$ takes the value 1; if q is assigned the value $1/2$ then, by Lemma 1, ϕ_1 and ϕ_2 and *a fortiori* their disjunction all take the value 1. This completes the proof that $\phi \vDash_{K3} \phi_1 \vee \phi_2$.

Were it the case that $\phi_1 \vee \phi_2 \vDash_{K3}$, so too would we have that $\phi \vDash_{K3}$, but by hypothesis this is not so. Hence, by the induction hypothesis, as $\phi_1 \vee \phi_2$ contains no more than m atoms that do not occur in ψ, an interpolant exists for $\phi_1 \vee \phi_2$ and ψ and the same formula suffices for ϕ and ψ. This completes the proof of the Interpolation Theorem for propositional $K3$. $\square$

References

Beall, Jc, Thomas Forster and Jeremy Seligman (2013), 'A Note on Freedom from Detachment in the Logic of Paradox', *Notre Dame Journal of Formal Logic*, **54**(1), 15–20.

Bendová, Kamila (2005), 'Interpolation and Three-valued Logics', *Reports on Mathematical Logic*, **39**, 127–31.

Hodges, Wilfrid (2001), *Logic: An Introduction to Elementary Logic*, second edition, London: Penguin.

Hunter, Geoffrey (1971), *Metalogic: An Introduction to the Metatheory of Standard First Order Logic*, Berkeley and Los Angeles: University of California Press.

Kleene, Stephen C. (1952), *Introduction to Metamathematics*, Amsterdam/ Groningen/New York: North-Holland/P. Noordhoff/D. van Nostrand.

Makinson, David (1973), *Topics in Modern Logic*, London: Methuen.

Priest, Graham (1979), 'The Logic of Paradox', *Journal of Philosophical Logic*, **8**(1), 219–41.

Renardel de Lavalette, Gerard R. (1981), 'The Interpolation Theorem in Fragments of Logics', *Indagationes Mathematicae*, **43**(1), 71–86.

Roorda, Dirk (1994), 'Interpolation in Fragments of Classical Linear Logic', *Journal of Symbolic Logic*, **59**(2), 419–44.

Schumm, George F. (1986), 'Some Failures of Interpolation in Modal Logic', *Notre Dame Journal of Formal Logic*, **27**(1), 108–10.

Takano, Mitio (1989), 'Interpolation Theorem in Many-valued Logics with Designated Values', *Kodai Mathematical Journal*, **12**(2), 125–31.

Urquhart, Alasdair (1993), 'Failure of Interpolation in Relevant Logics', *Journal of Philosophical Logic*, **22**(5), 449–79.

PETER MILNE
University of Stirling
peter.milne@stir.ac.uk