Details of Hunter's "Informal" Proof of Craig's Interpolation Theorem for P

Branden Fitelson

02/06/05

Hunter's proof of Craig's Interpolation theorem for P is a bit opaque. Here's a more detailed version of his proof, which I sketched in class on Friday. Since this is our first (non-trivial) metatheorem, it's worth doing a handout that proves it in some detail. We'll see similar kinds of proofs often in the course.

Theorem. Let A and B be formulas of P, such that (1) they share at least one propositional symbol in common, and (2) $\vDash_P A \supset B$. For any two such formulas of P, there exists a formula C (called the *P-interpolant* of the formulas A and B) such that (3) $\vDash_P A \supset C$, (4) $\vDash_P C \supset B$, and (5) C contains only propositional symbols that occur in both A and B (*i.e.*, only propositional symbols shared by A and B). [Intuitively, if $\vDash_P A \supset B$ (and A and B have *some* symbols in common!) it is always possible to reason from A to B *via* a formula C that has no propositional symbols not shared by A and B. This is sometimes called "linear reasoning" from A to B, since it takes no "detours" through "irrelevant" or "tangent" unshared propositional symbols.]

Proof. **Case 1**: There are zero propositional symbols occuring in A that do not also occur in B. That is, the set of propositional symbols in A is a subset of those in B. If we let $S(A)$ be the set of propositional symbols occurring in a formula A, then we can express this case as the case in which $S(A) \subseteq S(B)$. In this case, just let $C = A$. Then, obviously, (3) $\vDash_P A \supset C$, since $\vDash_P A \supset A$. And, since the assumption of the theorem is that $\vDash_P A \supset B$, we also know that $\vDash_P C \supset B$. All we need to show is that (5) C contains only propositional symbols that occur in both A and B. But, this follows from the fact that $C = A$, and the assumption of this Case, which is that $S(A) \subseteq S(B)$. Hence, $S(A) = S(C) = S(A) \cap S(B)$, which completes **Case 1**. □

Case 2: There are $n > 0$ propositional symbols occuring in A that do not also occur in B. That is, $\overline{\overline{S(A) - S(B)}} = n > 0$. We proceed by constructing a sequence of n interpolants $\langle C_1, \ldots, C_n \rangle$ in such a way that the last interpolant of the sequence C_n is such that (3) $\vDash_P A \supset C_n$, (4) $\vDash_P C_n \supset B$, and (5) C_n contains only propositional symbols that occur in both A and B (*i.e.*, only propositional symbols shared by A and B). This is an n-stage construction. Once you see how stage one works, the rest are just iterations.

- **Stage 1**: Our goal in **Stage 1** is to construct a formula C_1 such that (i) $\vDash_P A \supset C_1$, (ii) $\vDash_P C_1 \supset B$, and (iii) C_1 contains $n - 1$ propositional symbols that occur in A but not in B. Let p be some propositional symbol that occurs in A but not in B (there must be $n > 0$ of these, by the assumption of **Case 2**). Let q be some propositional symbol that occurs in both A and B (there must be at least one of these, by the assumption of the **Theorem**). Then, let A_1 be the formula you get when you replace all occurrences of p in A with $(q \supset q)$. And, let A_2 be the formula you get when you replace all occurrences of p in A with $\sim(q \supset q)$. It turns out that letting $C_1 = A_1 \vee A_2 = \sim A_1 \supset A_2$ does the trick. To prove this, we need to prove the following three things about $C_1 = A_1 \vee A_2 = \sim A_1 \supset A_2$:

  (i) $\vDash_P A \supset C_1$. That is, $\vDash_P A \supset (A_1 \vee A_2)$.

  Proof. Let I be an arbitrary interpretation. There are two cases: either (a) p is T on I, or (b) p is F on I. In case (a), A_1 must have the same truth-value as A, since the only difference between A and A_1 is that p gets replaced (in A) by something that is T on *all* interpretations, including I (the tautology $q \supset q$). But, in case (a), p was *already* T on I *before* it was replaced by $q \supset q$. So, the result (A_1) cannot have a different truth-value on I than A. In case (b), parallel reasoning shows that A_2 must have the same truth-value as A. Thus, on every interpretation I, either $A \supset A_1$ is T (if p is T on I) or $A \supset A_2$ is T (if p is F on I). Therefore, on every interpretation I, $A \supset (A_1 \vee A_2)$ is T on I. Technically, this is because $\vDash_P ((A \supset A_1) \vee (A \supset A_2)) \supset (A \supset (A_1 \vee A_2))$, which can be verified by truth-table reasoning. That completes the proof of (i). □

(*ii*) $\vDash_P C_1 \supset B$. That is, $\vDash_P (A_1 \vee A_2) \supset B$.

Proof. We will first prove that $\vDash_P A_1 \supset B$ and $\vDash_P A_2 \supset B$, from which (*ii*) follows by truth-table reasoning. Let's prove $\vDash_P A_1 \supset B$ first. Think of A as a truth-function of one argument: the truth-value of p, which I will call $\mathfrak{p}$. So, $A = f(\mathfrak{p})$. We can ignore the other arguments of A's truth-function, since $\mathfrak{p}$ is the only argument we're going to change. By definition, $A_1 = f(\mathsf{T})$. Now, it is clear from our definitions that $\vDash_P A \supset B$ is equivalent (in the metatheory) to:

For all I, and for all $\mathfrak{p} \in \{\mathsf{T}, \mathsf{F}\}$, either $f(\mathfrak{p}) = \mathsf{F}$ (A is F on I) or B is T on I.

Since $A_1 = f(\mathsf{T})$, we will have $\vDash_P A_1 \supset B$ just in case we have:

For all I, and for all $\mathfrak{p} \in \{\mathsf{T}\}$, either $f(\mathfrak{p}) = \mathsf{F}$ (A_1 is F on I) or B is T on I.

But, this is just a *special case* of the first claim, which quantifies over *all* $\mathfrak{p}$ (Note: the truth-value of B does not depend on $\mathfrak{p}$, since $p \notin S(B)$). This is a more rigorous way of making the point (which may not have been crystal clear in class) that $\vDash_P A \supset B$ entails $\vDash_P A_1 \supset B$.[1] A parallel quantificational meta-argument shows that $\vDash_P A \supset B$ entails $\vDash_P A_2 \supset B$, since $A_2 = f(\mathsf{F})$. It then follows from $\vDash_P A_1 \supset B$ and $\vDash_P A_2 \supset B$ that $\vDash_P (A_1 \vee A_2) \supset B$. Technically, this is because $\vDash_P ((A_1 \supset B) \wedge (A_2 \supset B)) \supset ((A_1 \vee A_2) \supset B)$, which can be verified by truth-table reasoning [note: $A \wedge B = \sim(A \supset \sim B)$]. And that completes the proof of (*ii*). ☐

(*iii*) C_1 has $n - 1$ propositional symbols that occur in A but not in B.

Proof. By the assumption of **Case 2**, the number of symbols in $S(A) - S(B)$ is $n > 0$. C_1 is constructed so that it contains *one less* such symbol (p is replaced by a function of q to form C_1). This completes the proof of (*iii*), and **Stage 1** of the construction. ☐

- **Stages 2 through** n: If we repeat the above construction, then we can form C_2, which removes *one more* symbol in $S(A) - S(B)$ from C_1, and which is such that $\vDash_P C_1 \supset C_2$ and $\vDash_P C_2 \supset B$. And, if we repeat this process $n - 2$ more times, then we will end-up with a chain of n such constructions $\vDash_P A \supset C_1$, $\vDash_P C_1 \supset C_2$, $\vDash_P C_2 \supset C_3 \ldots \vDash_P C_{n-1} \supset C_n$, $\vDash_P C_n \supset B$. Finally, because of the transitivity of material implication (*i.e.*, because if $\vDash_P X \supset Y$ and $\vDash_P Y \supset Z$, then $\vDash_P X \supset Z$), it will follow that (3) $\vDash_P A \supset C_n$, (4) $\vDash_P C_n \supset B$, and (5) C_n contains *zero* symbols in $S(A) - S(B)$ (*i.e.*, only propositional symbols that occur in both A and B), which is what we needed to show. ☐

This completes the ("informal") proof of **Case 2**, and with it the interpolation theorem. To do this more rigorously, we will need to prove it by *mathematical induction* on the cardinality of the set of propositional symbols $S(A) - S(B)$. We will do this soon. ☐

---

[1]**IMPORTANT NOTE**: The inference from $A \supset B$ to $A_1 \supset B$ (or to $A_2 \supset B$) is validity preserving, but *not* truth preserving! All we have shown here is that if $A \supset B$ is true on *all* interpretations, then so is $A_1 \supset B$. This does *not* imply that every interpretation on which $A \supset B$ is true is also an interpretation on which $A_1 \supset B$ is true. That is, we have *only* proved

If $\vDash_P A \supset B$, then $\vDash_P A_1 \supset B$.

We have *not* proven the following — *nor is the following true in the metatheory of* P!

$$\vDash_P (A \supset B) \supset (A_1 \supset B).$$

To see that this last meta-claim about P is *false*, consider the following concrete example. Let $A = (p'' \supset p''')$, and $B = (p'' \supset p')$. Then, $A \supset B$ $[(p'' \supset p''') \supset (p'' \supset p')]$ is *not* valid [it is F when p' is F, p'' is T, and p''' is T – check this!], but A and B otherwise satisfy the preconditions of the non-trivial case of Craig's theorem [they share one symbol (p'') and there is one symbol (p''') in $S(A) - S(B)$]. While $A \supset B$ is *not valid*, it is is true on *some* interpretations. For instance, $A \supset B$ is T whenever p'' is T, and p''' is F (check this!). But, $A_1 \supset B$ $[(p'' \supset (p'' \supset p'')) \supset (p'' \supset p')]$ is F on some of these interpretations. Specifically, $A_1 \supset B$ is F when p'' is T, p''' is F, and p' is F (check this!). So, this shows that the inference from $A \supset B$ to $A_1 \supset B$ is *not truth* preserving, even though it *is validity* preserving. A similar argument can be given to show that the inference from $A \supset B$ to $A_2 \supset B$ is *merely* validity preserving. As I mentioned on Friday, all truth preserving inferences are validity preserving. But, as this example explicitly shows, the converse of this entailment in the metatheory of P is *false*. So, there are ways of instantiating A and B such that $\nvDash_P (A \supset B) \supset (A_1 \supset B)$. But, these will always be cases in which both $\nvDash_P A \supset B$, and $\nvDash_P A_1 \supset B$.