

## int.1 Craig's Interpolation Theorem

mod:int:prf:  
 sec  
 mod:int:prf:  
 thm:interpol

**Theorem int.1** (Craig's Interpolation Theorem). *If  $\models \varphi \rightarrow \psi$ , then there is a sentence  $\chi$  such that  $\models \varphi \rightarrow \chi$  and  $\models \chi \rightarrow \psi$ , and every constant symbol, function symbol, and predicate symbol (other than  $=$ ) in  $\chi$  occurs both in  $\varphi$  and  $\psi$ . The sentence  $\chi$  is called an interpolant of  $\varphi$  and  $\psi$ .*

*Proof.* Suppose  $\mathcal{L}_1$  is the language of  $\varphi$  and  $\mathcal{L}_2$  is the language of  $\psi$ . Let  $\mathcal{L}_0 = \mathcal{L}_1 \cap \mathcal{L}_2$ . For each  $i \in \{0, 1, 2\}$ , let  $\mathcal{L}'_i$  be obtained from  $\mathcal{L}_i$  by adding the infinitely many new constant symbols  $c_0, c_1, c_2, \dots$ .

If  $\varphi$  is unsatisfiable,  $\exists x x \neq x$  is an interpolant. If  $\neg\psi$  is unsatisfiable (and hence  $\psi$  is valid),  $\exists x x = x$  is an interpolant. So we may assume also that both  $\varphi$  and  $\neg\psi$  are satisfiable.

In order to prove the contrapositive of the Interpolation Theorem, assume that there is no interpolant for  $\varphi$  and  $\psi$ . In other words, assume that  $\{\varphi\}$  and  $\{\neg\psi\}$  are inseparable in  $\mathcal{L}_0$ .

Our goal is to extend the pair  $(\{\varphi\}, \{\neg\psi\})$  to a maximally inseparable pair  $(\Gamma^*, \Delta^*)$ . Let  $\varphi_0, \varphi_1, \varphi_2, \dots$  enumerate the sentences of  $\mathcal{L}_1$ , and  $\psi_0, \psi_1, \psi_2, \dots$  enumerate the sentences of  $\mathcal{L}_2$ . We define two increasing sequences of sets of sentences  $(\Gamma_n, \Delta_n)$ , for  $n \geq 0$ , as follows. Put  $\Gamma_0 = \{\varphi\}$  and  $\Delta_0 = \{\neg\psi\}$ . Assuming  $(\Gamma_n, \Delta_n)$  are already defined, define  $\Gamma_{n+1}$  and  $\Delta_{n+1}$  by:

1. If  $\Gamma_n \cup \{\varphi_n\}$  and  $\Delta_n$  are inseparable in  $\mathcal{L}'_0$ , put  $\varphi_n$  in  $\Gamma_{n+1}$ . Moreover, if  $\varphi_n$  is an existential formula  $\exists x \sigma$  then pick a new constant symbol  $c$  not occurring in  $\Gamma_n, \Delta_n, \varphi_n$  or  $\psi_n$ , and put  $\sigma[c/x]$  in  $\Gamma_{n+1}$ .
2. If  $\Gamma_{n+1}$  and  $\Delta_n \cup \{\psi_n\}$  are inseparable in  $\mathcal{L}'_0$ , put  $\psi_n$  in  $\Delta_{n+1}$ . Moreover, if  $\psi_n$  is an existential formula  $\exists x \sigma$ , then pick a new constant symbol  $c$  not occurring in  $\Gamma_{n+1}, \Delta_n, \varphi_n$  or  $\psi_n$ , and put  $\sigma[c/x]$  in  $\Delta_{n+1}$ .

Finally, define:

$$\Gamma^* = \bigcup_{n \geq 0} \Gamma_n, \quad \Delta^* = \bigcup_{n \geq 0} \Delta_n.$$

By simultaneous induction on  $n$  we can now prove:

mod:int:prf:  
 part-a  
 mod:int:prf:  
 part-b

1.  $\Gamma_n$  and  $\Delta_n$  are inseparable in  $\mathcal{L}'_0$ ;
2.  $\Gamma_{n+1}$  and  $\Delta_n$  are inseparable in  $\mathcal{L}'_0$ .

The basis for (1) is given by ???. For part (2), we need to distinguish three cases:

1. If  $\Gamma_0 \cup \{\varphi_0\}$  and  $\Delta_0$  are separable, then  $\Gamma_1 = \Gamma_0$  and (2) is just (1);
2. If  $\Gamma_1 = \Gamma_0 \cup \{\varphi_0\}$ , then  $\Gamma_1$  and  $\Delta_0$  are inseparable by construction.

3. It remains to consider the case where  $\varphi_0$  is existential, so that  $\Gamma_1 = \Gamma_0 \cup \{\exists x \sigma, \sigma[c/x]\}$ . By construction,  $\Gamma_0 \cup \{\exists x \sigma\}$  and  $\Delta_0$  are inseparable, so that by ?? also  $\Gamma_0 \cup \{\exists x \sigma, \sigma[c/x]\}$  and  $\Delta_0$  are inseparable.

This completes the basis of the induction for (1) and (2) above. Now for the inductive step. For (1), if  $\Delta_{n+1} = \Delta_n \cup \{\psi_n\}$  then  $\Gamma_{n+1}$  and  $\Delta_{n+1}$  are inseparable by construction (even when  $\psi_n$  is existential, by ??); if  $\Delta_{n+1} = \Delta_n$  (because  $\Gamma_{n+1}$  and  $\Delta_n \cup \{\psi_n\}$  are separable), then we use the induction hypothesis on (2). For the inductive step for (2), if  $\Gamma_{n+2} = \Gamma_{n+1} \cup \{\varphi_{n+1}\}$  then  $\Gamma_{n+2}$  and  $\Delta_{n+1}$  are inseparable by construction (even when  $\varphi_{n+1}$  is existential, by ??); and if  $\Gamma_{n+2} = \Gamma_{n+1}$  then we use the inductive case for (1) just proved. This concludes the induction on (1) and (2).

It follows that  $\Gamma^*$  and  $\Delta^*$  are inseparable; if not, by compactness, there is  $n \geq 0$  that separates  $\Gamma_n$  and  $\Delta_n$ , against (1). In particular,  $\Gamma^*$  and  $\Delta^*$  are consistent: for if the former or the latter is inconsistent, then they are separated by  $\exists x x \neq x$  or  $\forall x x = x$ , respectively.

We now show that  $\Gamma^*$  is maximally consistent in  $\mathcal{L}'_1$  and likewise  $\Delta^*$  in  $\mathcal{L}'_2$ . For the former, suppose that  $\varphi_n \notin \Gamma^*$  and  $\neg\varphi_n \notin \Gamma^*$ , for some  $n \geq 0$ . If  $\varphi_n \notin \Gamma^*$  then  $\Gamma_n \cup \{\varphi_n\}$  is separable from  $\Delta_n$ , and so there is  $\chi \in \mathcal{L}'_0$  such that both:

$$\Gamma^* \models \varphi_n \rightarrow \chi, \quad \Delta^* \models \neg\chi.$$

Likewise, if  $\neg\varphi_n \notin \Gamma^*$ , there is  $\chi' \in \mathcal{L}'_0$  such that both:

$$\Gamma^* \models \neg\varphi_n \rightarrow \chi', \quad \Delta^* \models \neg\chi'.$$

By propositional logic,  $\Gamma^* \models \chi \vee \chi'$  and  $\Delta^* \models \neg(\chi \vee \chi')$ , so  $\chi \vee \chi'$  separates  $\Gamma^*$  and  $\Delta^*$ . A similar argument establishes that  $\Delta^*$  is maximal.

Finally, we show that  $\Gamma^* \cap \Delta^*$  is maximally consistent in  $\mathcal{L}'_0$ . It is obviously consistent, since it is the intersection of consistent sets. To show maximality, let  $\sigma \in \mathcal{L}'_0$ . Now,  $\Gamma^*$  is maximal in  $\mathcal{L}'_1 \supseteq \mathcal{L}'_0$ , and similarly  $\Delta^*$  is maximal in  $\mathcal{L}'_2 \supseteq \mathcal{L}'_0$ . It follows that either  $\sigma \in \Gamma^*$  or  $\neg\sigma \in \Gamma^*$ , and either  $\sigma \in \Delta^*$  or  $\neg\sigma \in \Delta^*$ . If  $\sigma \in \Gamma^*$  and  $\neg\sigma \in \Delta^*$  then  $\sigma$  would separate  $\Gamma^*$  and  $\Delta^*$ ; and if  $\neg\sigma \in \Gamma^*$  and  $\sigma \in \Delta^*$  then  $\Gamma^*$  and  $\Delta^*$  would be separated by  $\neg\sigma$ . Hence, either  $\sigma \in \Gamma^* \cap \Delta^*$  or  $\neg\sigma \in \Gamma^* \cap \Delta^*$ , and  $\Gamma^* \cap \Delta^*$  is maximal.

Since  $\Gamma^*$  is maximally consistent, it has a model  $\mathfrak{M}'_1$  whose domain  $|\mathfrak{M}'_1|$  comprises all and only the elements  $c^{\mathfrak{M}'_1}$  interpreting the **constant symbols**—just like in the proof of the completeness theorem (?). Similarly,  $\Delta^*$  has a model  $\mathfrak{M}'_2$  whose domain  $|\mathfrak{M}'_2|$  is given by the interpretations  $c^{\mathfrak{M}'_2}$  of the **constant symbols**.

Let  $\mathfrak{M}_1$  be obtained from  $\mathfrak{M}'_1$  by dropping interpretations for **constant symbols**, **function symbols**, and **predicate symbols** in  $\mathcal{L}'_1 \setminus \mathcal{L}'_0$ , and similarly for  $\mathfrak{M}_2$ . Then the map  $h: M_1 \rightarrow M_2$  defined by  $h(c^{\mathfrak{M}'_1}) = c^{\mathfrak{M}'_2}$  is an isomorphism in  $\mathcal{L}'_0$ , because  $\Gamma^* \cap \Delta^*$  is maximally consistent in  $\mathcal{L}'_0$ , as shown. This follows because any  $\mathcal{L}'_0$ -sentence either belongs to both  $\Gamma^*$  and  $\Delta^*$ , or to neither: so  $c^{\mathfrak{M}'_1} \in P^{\mathfrak{M}'_1}$  if and only if  $P(c) \in \Gamma^*$  if and only if  $P(c) \in \Delta^*$  if and only if

$c^{\mathfrak{M}'_2} \in P^{\mathfrak{M}'_2}$ . The other conditions satisfied by isomorphisms can be established similarly.

Let us now define a model  $\mathfrak{M}$  for the language  $\mathcal{L}_1 \cup \mathcal{L}_2$  as follows:

1. The domain  $|\mathfrak{M}|$  is just  $|\mathfrak{M}_2|$ , i.e., the set of all elements  $c^{\mathfrak{M}'_2}$ ;
2. If a predicate symbol  $P$  is in  $\mathcal{L}_2 \setminus \mathcal{L}_1$  then  $P^{\mathfrak{M}} = P^{\mathfrak{M}'_2}$ ;
3. If a predicate  $P$  is in  $\mathcal{L}_1 \setminus \mathcal{L}_2$  then  $P^{\mathfrak{M}} = h(P^{\mathfrak{M}'_2})$ , i.e.,  $\langle c_1^{\mathfrak{M}'_2}, \dots, c_n^{\mathfrak{M}'_2} \rangle \in P^{\mathfrak{M}}$  if and only if  $\langle c_1^{\mathfrak{M}'_1}, \dots, c_n^{\mathfrak{M}'_1} \rangle \in P^{\mathfrak{M}'_1}$ .
4. If a predicate symbol  $P$  is in  $\mathcal{L}_0$  then  $P^{\mathfrak{M}} = P^{\mathfrak{M}'_2} = h(P^{\mathfrak{M}'_1})$ .
5. Function symbols of  $\mathcal{L}_1 \cup \mathcal{L}_2$ , including constant symbols, are handled similarly.

Finally, one shows by induction on formulas that  $\mathfrak{M}$  agrees with  $\mathfrak{M}'_1$  on all formulas of  $\mathcal{L}'_1$  and with  $\mathfrak{M}'_2$  on all formulas of  $\mathcal{L}'_2$ . In particular,  $\mathfrak{M} \models \Gamma^* \cup \Delta^*$ , whence  $\mathfrak{M} \models \varphi$  and  $\mathfrak{M} \models \neg\psi$ , and  $\not\models \varphi \rightarrow \psi$ . This concludes the proof of Craig's Interpolation Theorem.  $\square$

## Photo Credits

## Bibliography