Chapter udf

The Interpolation Theorem

int.1  Introduction

The interpolation theorem is the following result: Suppose $\models \varphi \rightarrow \psi$. Then there is a sentence $\chi$ such that $\models \varphi \rightarrow \chi$ and $\models \chi \rightarrow \psi$. Moreover, 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 $\psi$.

The interpolation theorem is interesting in its own right, but its main importance lies in the fact that it can be used to prove results about definability in a theory, and the conditions under which combining two consistent theories results in a consistent theory. The first result is known as the Beth definability theorem; the second, Robinson’s joint consistency theorem.

int.2  Separation of Sentences

A bit of groundwork is needed before we can proceed with the proof of the interpolation theorem. An interpolant for $\varphi$ and $\psi$ is a sentence $\chi$ such that $\varphi \models \chi$ and $\chi \models \psi$. By contraposition, the latter is true iff $\neg \psi \models \neg \chi$. A sentence $\chi$ with this property is said to separate $\varphi$ and $\neg \psi$. So finding an interpolant for $\varphi$ and $\psi$ amounts to finding a sentence that separates $\varphi$ and $\neg \psi$. As so often, it will be useful to consider a generalization: a sentence that separates two sets of sentences.

Definition int.1. A sentence $\chi$ separates sets of sentences $\Gamma$ and $\Delta$ if and only if $\Gamma \models \chi$ and $\Delta \models \neg \chi$. If no such sentence exists, then $\Gamma$ and $\Delta$ are inseparable.

The inclusion relations between the classes of models of $\Gamma$, $\Delta$ and $\chi$ are represented below:

Lemma int.2. Suppose $\mathcal{L}_0$ is the language containing every constant symbol, function symbol and predicate symbol (other than $=$) that occurs in both $\Gamma$ and $\Delta$, and let $\mathcal{L}_0'$ be obtained by the addition of infinitely many new constant
symbols $c_n$ for $n \geq 0$. Then if $\Gamma$ and $\Delta$ are inseparable in $\mathcal{L}_0$, they are also inseparable in $\mathcal{L}'_0$.

**Proof.** We proceed indirectly: suppose by way of contradiction that $\Gamma$ and $\Delta$ are separated in $\mathcal{L}'_0$. Then $\Gamma \models \chi[c/x]$ and $\Delta \models \neg \chi[c/x]$ for some $\chi \in \mathcal{L}_0$ (where $c$ is a new constant symbol—the case where $\chi$ contains more than one such new constant symbol is similar). By compactness, there are finite subsets $\Gamma_0$ of $\Gamma$ and $\Delta_0$ of $\Delta$ such that $\Gamma_0 \models \chi[c/x]$ and $\Delta_0 \models \neg \chi[c/x]$. Let $\gamma$ be the conjunction of all formulas in $\Gamma_0$ and $\delta$ the conjunction of all formulas in $\Delta_0$. Then

$$
\gamma \models \chi[c/x], \quad \delta \models \neg \chi[c/x].
$$

From the former, by Generalization, we have $\gamma \models \forall x \chi$, and from the latter by contraposition, $\chi[c/x] \models \neg \delta$, whence also $\forall x \chi \models \neg \delta$. Contraposition again gives $\delta \models \neg \forall x \chi$. By monotonicity,

$$
\Gamma \models \forall x \chi, \quad \Delta \models \neg \forall x \chi,
$$

so that $\forall x \chi$ separates $\Gamma$ and $\Delta$ in $\mathcal{L}_0$.

**Lemma int.3.** Suppose that $\Gamma \cup \{\exists x \sigma\}$ and $\Delta$ are inseparable, and $c$ is a new constant symbol not in $\Gamma$, $\Delta$, or $\sigma$. Then $\Gamma \cup \{\exists x \sigma, \sigma[c/x]\}$ and $\Delta$ are also inseparable.

**Proof.** Suppose for contradiction that $\chi$ separates $\Gamma \cup \{\exists x \sigma, \sigma[c/x]\}$ and $\Delta$, while at the same time $\Gamma \cup \{\exists x \sigma\}$ and $\Delta$ are inseparable. We distinguish two cases:

1. $c$ does not occur in $\chi$: in this case $\Gamma \cup \{\exists x \sigma, \neg \chi\}$ is satisfiable (otherwise $\chi$ separates $\Gamma \cup \{\exists x \sigma\}$ and $\Delta$). It remains so if $\sigma[c/x]$ is added, so $\chi$ does not separate $\Gamma \cup \{\exists x \sigma, \sigma[c/x]\}$ and $\Delta$ after all.

2. $c$ does occur in $\chi$ so that $\chi$ has the form $\chi[c/x]$. Then we have that

$$
\Gamma \cup \{\exists x \sigma, \sigma[c/x]\} \models \chi[c/x],
$$

whence $\Gamma, \exists x \sigma \models \forall x (\sigma \rightarrow \chi)$ by the Deduction Theorem and Generalization, and finally $\Gamma \cup \{\exists x \sigma\} \models \exists x \chi$. On the other hand, $\Delta \models \neg \chi[c/x]$ and hence by Generalization $\Delta \models \neg \exists x \chi$. So $\Gamma \cup \{\exists x \sigma\}$ and $\Delta$ are separable, a contradiction. 

---

*interpolation rev: ad37848 (2024-05-01) by OLP / CC–BY*
Craig’s Interpolation Theorem

Theorem int.4 (Craig’s Interpolation Theorem). If $\varphi \rightarrow \psi$, 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 $L_1$ is the language of $\varphi$ and $L_2$ is the language of $\psi$. Let $L_0 = L_1 \cap L_2$. For each $i \in \{0, 1, 2\}$, let $L^*_i$ be obtained from $L_i$ by adding the infinitely many new constant symbols $c_0, c_1, c_2, \ldots$

If $\varphi$ is unsatisfiable ($\models \neg \psi$ is valid), $\exists x x \neq x$ 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 $L_0$.

Our goal is to extend the pair $\{\varphi\}, \{\neg \psi\}$ to a maximally inseparable pair $(\Gamma^*, \Delta^*)$. Let $\varphi_0, \varphi_1, \varphi_2, \ldots$ enumerate the sentences of $L_1$, and $\psi_0, \psi_1, \psi_2, \ldots$ enumerate the sentences of $L_2$. 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 $L'_0$, 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 $L'_0$, 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:

1. $\Gamma_n$ and $\Delta_n$ are inseparable in $L'_0$;

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

The basis for (1) is given by Lemma int.2. 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.
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 Lemma int.3); 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 Lemma int.3); 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 \( \mathcal{L}'_2 \) 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 \( \chi \in \mathcal{L}'_2 \) such that both:

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

Likewise, if \( \neg \varphi_n \notin \Gamma^* \), there is \( \chi' \in \mathcal{L}'_2 \) such that both:

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

By propositional logic, \( \Gamma^* \models \chi \lor \chi' \) and \( \Delta^* \models \neg (\chi \lor \chi') \), so \( \chi \lor \chi' \) 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 \( \sigma \in \Gamma^* \cap \Delta^* \) then \( \sigma \) would separate \( \Gamma^* \) and \( \Delta^* \); and if \( \neg \sigma \in \Gamma^* \cap \Delta^* \) 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 \( \mathcal{M}'_1 \) whose domain \( |\mathcal{M}'_1| \) comprises all and only the elements \( c_{\mathcal{M}'_1} \) interpreting the constant symbols—just like in the proof of the completeness theorem (??). Similarly, \( \Delta^* \) has a model \( \mathcal{M}'_2 \) whose domain \( |\mathcal{M}'_2| \) is given by the interpretations \( c_{\mathcal{M}'_2} \) of the constant symbols.

Let \( \mathcal{M}_1 \) be obtained from \( \mathcal{M}'_1 \) by dropping interpretations for constant symbols, function symbols, and predicate symbols in \( L_1 \setminus L'_0 \), and similarly for \( \mathcal{M}_2 \). Then the map \( h: M_1 \rightarrow M_2 \) defined by \( h(c_{\mathcal{M}'_1}) = c_{\mathcal{M}'_2} \) 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
The Interpolation Theorem.

whence ML formulas of interpolation

Suppose Definition int.5.

P bols

Suppose Definition int.6.

P bol

implicitly defines in this way—whenever it

P of

ing completely—a relation. A theory may also be such that the interpretation

entails) a formalized explicit definition, i.e.,

the

n terms of \( \chi \) variables which does not involve \( c \).

One important application of the interpolation theorem is Beth’s definability

theorem. To define an

n-place predicate symbol \( P \) explicitly defines \( P \) if and only if \( \chi(x_1, \ldots, x_n) \leftrightarrow \chi(x_1, \ldots, x_n) \).

But an explicit definition is only one way of defining—in the sense of determining completely—a relation. A theory may also be such that the interpretation of \( P \) is fixed by the interpretation of the rest of the language in any model. The definability theorem states that whenever a theory fixes the interpretation of \( P \) in this way—whenever it implicitly defines \( P \)—then it also explicitly defines it.

Definition int.5. Suppose \( L \) is a language not containing the predicate symbol \( P \). A set \( \Sigma(P) \) of sentences of \( L \cup \{ P \} \) explicitly defines \( P \) if and only if there is a formula \( \chi(x_1, \ldots, x_n) \) of \( L \) such that

\[ \Sigma(P) \models \forall x_1 \ldots \forall x_n (P(x_1, \ldots, x_n) \leftrightarrow \chi(x_1, \ldots, x_n)). \]

Definition int.6. Suppose \( L \) is a language not containing the predicate symbols \( P \) and \( P' \). A set \( \Sigma(P) \) of sentences of \( L \cup \{ P \} \) implicitly defines \( P \) if and only if

\[ \Sigma(P) \cup \Sigma(P') \models \forall x_1 \ldots \forall x_n (P(x_1, \ldots, x_n) \leftrightarrow P'(x_1, \ldots, x_n)). \]
where \( \Sigma(P') \) is the result of uniformly replacing \( P \) with \( P' \) in \( \Sigma(P) \).

In other words, for any model \( \mathfrak{M} \) and \( R, R' \subseteq |\mathfrak{M}|^n \), if both \( (\mathfrak{M}, R) \models \Sigma(P) \) and \( (\mathfrak{M}, R') \models \Sigma(P') \), then \( R = R' \); where \( (\mathfrak{M}, R) \) is the structure \( \mathfrak{M}' \) for the expansion of \( \mathcal{L} \) to \( \mathcal{L} \cup \{P\} \) such that \( P^{\mathfrak{M}'} = R \), and similarly for \( (\mathfrak{M}, R') \).

**Theorem int.7 (Beth Definability Theorem).** A set \( \Sigma(P) \) of \( \mathcal{L} \cup \{P\} \)-formulas implicitly defines \( P \) if and only \( \Sigma(P) \) explicitly defines \( P \).

**Proof.** If \( \Sigma(P) \) explicitly defines \( P \) then both

\[
\Sigma(P) \models \forall x_1 \ldots \forall x_n (P(x_1, \ldots, x_n) \leftrightarrow \chi(x_1, \ldots, x_n))
\]

\[
\Sigma(P') \models \forall x_1 \ldots \forall x_n (P'(x_1, \ldots, x_n) \leftrightarrow \chi(x_1, \ldots, x_n))
\]

and the conclusion follows. For the converse: assume that \( \Sigma(P) \) implicitly defines \( P \). First, we add constant symbols \( c_1, \ldots, c_n \) to \( \mathcal{L} \). Then

\[
\Sigma(P) \cup \Sigma(P') \models P(c_1, \ldots, c_n) \rightarrow P'(c_1, \ldots, c_n).
\]

By compactness, there are finite sets \( \Delta_0 \subseteq \Sigma(P) \) and \( \Delta_1 \subseteq \Sigma(P') \) such that

\[
\Delta_0 \cup \Delta_1 \models P(c_1, \ldots, c_n) \rightarrow P'(c_1, \ldots, c_n).
\]

Let \( \theta(P) \) be the conjunction of all sentences \( \varphi(P) \) such that either \( \varphi(P) \in \Delta_0 \) or \( \varphi(P') \in \Delta_1 \) and let \( \theta(P') \) be the conjunction of all sentences \( \varphi(P') \) such that either \( \varphi(P) \in \Delta_0 \) or \( \varphi(P) \in \Delta_1 \). Then \( \theta(P) \land \theta(P') \models P(c_1, \ldots, c_n) \rightarrow P'(c_1, \ldots, c_n) \). We can re-arrange this so that each predicate symbol occurs on one side of \( \models \):

\[
\theta(P) \land P(c_1, \ldots, c_n) \models \theta(P') \rightarrow P'(c_1, \ldots, c_n).
\]

By Craig’s Interpolation Theorem there is a sentence \( \chi(c_1, \ldots, c_n) \) not containing \( P \) or \( P' \) such that:

\[
\theta(P) \land P(c_1, \ldots, c_n) \models \chi(c_1, \ldots, c_n) ; \quad \chi(c_1, \ldots, c_n) \models \theta(P') \rightarrow P'(c_1, \ldots, c_n).
\]

From the former of these two entailments we have: \( \theta(P) \models P(c_1, \ldots, c_n) \rightarrow \chi(c_1, \ldots, c_n) \). And from the latter, since an \( \mathcal{L} \cup \{P\} \)-model \( (\mathfrak{M}, R) \models \varphi(P) \) if and only if the corresponding \( \mathcal{L} \cup \{P'\} \)-model \( (\mathfrak{M}, R) \models \varphi(P) \), we have \( \chi(c_1, \ldots, c_n) \models \theta(P) \rightarrow P(c_1, \ldots, c_n) \), from which:

\[
\theta(P) \models \chi(c_1, \ldots, c_n) \rightarrow P(c_1, \ldots, c_n).
\]

Putting the two together, \( \theta(P) \models P(c_1, \ldots, c_n) \leftrightarrow \chi(c_1, \ldots, c_n) \), and by monotonicity and generalization also

\[
\Sigma(P) \models \forall x_1 \ldots \forall x_n (P(x_1, \ldots, x_n) \leftrightarrow \chi(x_1, \ldots, x_n)).
\]

\(\square\)
Photo Credits
Bibliography