int.1 The Definability Theorem

One important application of the interpolation theorem is Beth’s definability theorem. To define an \( n \)-place relation \( R \) we can give a formula \( \chi \) with \( n \) free variables which does not involve \( R \). This would be an explicit definition of \( R \) in terms of \( \chi \). We can then say also that a theory \( \Sigma(P) \) in a language containing the \( n \)-place predicate symbol \( P \) explicitly defines \( P \) if it contains (or at least entails) a formalized explicit definition, i.e.,

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

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.1.** Suppose \( \mathcal{L} \) is a language not containing the predicate symbol \( P \). A set \( \Sigma(P) \) of sentences of \( \mathcal{L} \cup \{P\} \) explicitly defines \( P \) if and only if there is a formula \( \chi(x_1, \ldots, x_n) \) of \( \mathcal{L} \) 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.2.** Suppose \( \mathcal{L} \) is a language not containing the predicate symbols \( P \) and \( P' \). A set \( \Sigma(P) \) of sentences of \( \mathcal{L} \cup \{P\} \) 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 \( \mathfrak{M}^R \) for the expansion of \( \mathcal{L} \) to \( \mathcal{L} \cup \{P\} \) such that \( P^{\mathfrak{M}^R} = R \), and similarly for \( (\mathfrak{M}, R') \).

**Theorem int.3 (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

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

\[\Box\]

Photo Credits

Bibliography