AUTOMORPHISM AND DEFINABILITY

Some notes on Thursday’s lecture. We now focus on first-order logic.

- We defined formulas and sentences for a language \( \mathcal{L} \).
- We defined models and gave a recursive definition for when a sentence \( \varphi \) is true in a model \( M \), written \( M \models \varphi \). Models are sometimes called \( \mathcal{L} \)-structures to emphasize exactly which language is at stake.
- For any \( \mathcal{L} \)-structure \( M \), recall that \( Th(M) \), the theory of \( M \), is the set of all \( \mathcal{L} \)-sentences \( \varphi \) such that \( M \models \varphi \). We defined \( M \equiv N \) to mean \( Th(M) = Th(N) \) (so this is an equivalence relation on the class of \( \mathcal{L} \)-structures).
- An automorphism of a model \( M \) is a bijection \( f : \text{dom}(M) \to \text{dom}(M) \) which preserves relations, functions, and constants. More precisely, we require that:
  1. for each constant \( c \in \mathcal{L}, f(c^M) = c^M \).
  2. for each \( n \in \mathbb{N} \), each \( n \)-ary relation symbol \( R \) of \( \mathcal{L} \), and all \( a_1, \ldots, a_n \in \text{dom}(M) \), \( R^M(a_1, \ldots, a_n) \iff R^M(f(a_1), \ldots, f(a_n)) \).
  3. for each \( n \in \mathbb{N} \), each \( n \)-ary function symbol \( F \) of \( \mathcal{L} \), and all \( a_1, \ldots, a_n \in \text{dom}(M) \), \( F^M(a_1, \ldots, a_n) = F^M(f(a_1), \ldots, f(a_n)) \).
- Example. Let \( \mathcal{L} = \{ R \} \), a single binary relation. Let \( M \) have domain \( \mathbb{Z} \) and suppose \( R \) is interpreted as \( \{(n, -n) : n \in \mathbb{N} \} \). By our definition, the map given by \( n \mapsto -n \) is an automorphism of \( M \), whereas the map given by \( n \mapsto n + 1 \) is not. However, if we let \( M' \) have domain \( \mathbb{Z} \) but interpret \( R \) as \( < \), then \( n \mapsto n + 1 \) is an automorphism of \( M' \) whereas \( n \mapsto -n \) is not.
- Let \( A \subseteq (\text{dom}(M))^k \) be a set of \( k \)-tuples of elements of \( M \). Recall that \( A \) is definable if there exists some formula \( \psi(x_1, \ldots, x_k) \) such that for any \( a_1, \ldots, a_k \in \text{dom}(M) \), \( M \models \psi(a_1, \ldots, a_k) \) if and only if \( a_1, \ldots, a_k \in A \).

We will prove (after the exam, but in the meantime you are free to appeal to this fact) that

**Theorem 1.** If \( A \subseteq (\text{dom}(M))^k \) is a definable set, then any automorphism \( f \) of \( M \) preserves \( A \) setwise, i.e. \( f(A) = A \).

We discussed in class that in some sense, “most sets are not definable.” However, really our only technique so far for showing a set is not definable is to apply this theorem. On the other hand, to show a set is definable is easier: simply write down a defining formula.

**Examples.** Let \( \mathcal{L} = \{ R \} \) where \( R \) is a binary relation and \( S \) is a unary function.

Example 1: Let \( M \) be the model whose domain is \( \mathbb{Z} \), \( R \) is interpreted as \(<\). Show that \( \mathbb{N} \) is not a definable set in \( M \). [Use the contrapositive of the theorem: give an automorphism which does not preserve it.]

Example 2: Let \( M \) be the model whose domain is \( \mathbb{N} \), \( R \) is \(<\). Show that the graph of the successor function is a definable set. [Write down a formula.]

**WARNING.** Beware the converse: Just because a set is preserved by automorphisms need not mean it is definable. For instance, \( \langle \mathbb{N}, S \rangle \) has no nontrivial automorphisms (why?), but many subsets of its domain are not definable (why?).