Lecture 20: Expressiveness of FOL - I

1 Expressiveness of First Order Logic

We ask two related types of questions about expressiveness:

- *Definability*: What things can be defined in the language? Here we want to know what structures, relations and other mathematical objects can be defined using first order logic.
- *Distinguishability*: What notions of 'things being different' can the language capture? Here we want to know when two mathematical structures are the same or similar with respect to first order logic.

Note that distinguishability and equivalence are two sides of the same coin. In this lecture we will discuss equivalence, and the next lecture will focus on definability.

2 Homomorphisms and Isomorphisms

In mathematics, we use functions or mappings to relate sets to each other. For example, the square function relates $\mathbb{R}$ and $\mathbb{R}^+$ because the square of every real number is positive. However, we know that $\mathbb{R}$ is a group under addition, while $\mathbb{R}^+$ is not, so some of the structure of the reals is lost by this mapping.

As another example, consider the functions $f_1 : \mathbb{Z} \to \mathbb{Z}$, $f_2 : \mathbb{Z} \to \mathbb{Z}$ and $f_3 : \mathbb{Z} \to \mathbb{Z}$ defined on the integers as follows: $f_1(m) = m - 1$, $f_2(m) = 2m$, $f_3(m) = m$ if m is even and $-m$ otherwise. Then f_1 and f_2 preserve the less-than order relation $<$ on $\mathbb{Z}$, because if $x < y$ then $x - 1 < y - 1$ and $2x < 2y$. Geometrically, f_1 defines a translation of the integers, while f_2 scales (stretches) the integer line. It is intuitively clear that both rigid translation and scaling should not disturb the order of elements. Further, f_2 also preserves the addition operation, because $2(x + y) = 2x + 2y$, but f_1 does not preserve addition. So scaling preserves addition, but translation does not. Finally, f_3 preserves neither the order nor the additive structure of the integers. Thus we see that some mappings preserve certain structures while others do not.

We now formalize what it means for a mapping to 'preserve structure'.

Definition 1 (Homomorphism)**.** *Let A and B be structures over the same vocabulary. A mapping $h : D^A \to D^B$ is a* homomorphism *if*

1. For each k-ary function symbol f and each $(a_1, \ldots, a_k) \in (D^A)^k$,

$$h(f^A(a_1, \ldots, a_k)) = f^B(h(a_1), \ldots, h(a_k))$$

2. For each k-ary relation symbol P and each $(a_1, \ldots, a_k) \in (D^A)^k$,

$$(a_1, \ldots, a_k) \in P^A \textit{ iff } (h(a_1), \ldots, h(a_k)) \in P^B$$

Example 1. Let $A = (N, +)$ (natural numbers with addition) and $B = (\{0, 1\}, +_b)$ (one-bit numbers with boolean addition). Let $h : N \to \{0, 1\}$ be defined as

$$h(n) = \begin{cases} 0, & n \text{ is even} \\ 1, & n \text{ is odd} \end{cases}$$

Then h is a homomorphism from A to B.
Proof: We have to show that for all m, n in N, $h(m + n) = h(m) +_b h(n)$.
Case 1: $h(odd + even) = h(odd) = 1$ and $h(odd) +_b h(even) = 1 +_b 0 = 1$.
Case 2: $h(odd + odd) = h(even) = 0$ and $h(odd) +_b h(odd) = 1 +_b 1 = 0$.
Case 3: $h(even + even) = h(even) = 0$ and $h(even) +_b h(even) = 0 +_b 0 = 0$.

Example 2. Let $A = (\mathbb{C}, \cdot)$ (complex numbers with multiplication) and $B = (\mathbb{R}, \cdot)$ (real numbers with multiplication). Then $h : \mathbb{C} \to \mathbb{R}$, defined as $h(z) = |z|$, is a homomorphism from A to B. (here $|z|$ denotes the modulus of complex number z).
Proof: We have $|z_1 \cdot z_2| = |z_1| \cdot |z_2|$ by the definition of the modulus.

We can view a homomorphism as an abstraction, that selectively preserves information and throws away the rest. For example, in Example 1, h preserves parity information (i.e., whether a number is even or odd) but does not preserve other properties such as size or primality. Similarly, the modulus of a complex number only retains distance information, but abstracts away the phase (angle with x-axis) of the complex number. This loss of information occurs because, in general, a homomorphism is not reversible. If, however, the homomorphism is also a bijection then all structural information is preserved. This is called an *isomorphism*

Definition 2 (Isomorphism)**.** *A homomorphism* $h : D^A \to D^B$ *is an* isomorphism *if it is one-to-one and onto. If h is an isomorphism from A to B, then we say that A is isomorphic to B, denoted by* $A \simeq B$. *Note that* $\simeq$ *is an equivalence relation called* isomorphic equivalence.

An isomorphism from A to B is essentially a renaming of elements in D^A to elements in D^B, while preserving all the structure. To show that h is an isomorphism requires four distinct steps:

1. Show that h is a function, i.e., h maps each element of D^A to a unique element of D^B (this is usually trivial).
2. Show that h is one-to-one: $\forall x \forall y \; h(x) = h(y) \Rightarrow x = y$.

3. Show that h is onto: $\forall y \exists x\ h(x) = y$
4. Show that h is a homomorphism.

Example 3. $(\mathbb{R}^+, \cdot)$ and $(\mathbb{R}, +)$ are isomorphic. The isomorphism is given by $h : \mathbb{R}^+ \to \mathbb{R}$, defined as $h(x) = ln(x)$.
Proof: Since the natural log of any positive real is unique, h is a function. Since the logarithmic function $ln(x)$ is strictly increasing, we have $x \neq y \Rightarrow ln(x) \neq ln(y)$. So h is one-to-one. Given any $x \in \mathbb{R}$, $ln(e^x) = x$ and $e^x \in \mathbb{R}^+$. So h is onto. Finally, $ln(x \cdot y) = ln(x) + ln(y)$ is a fundamental property of logarithms. So h is a homomorphism. Thus, h is an isomorphism.

Example 4. $(\mathbb{C}, +)$ and $(\mathbb{R}^2, +)$ are isomorphic. The isomorphism is given by $(a + ib) \mapsto (a, b)$.

While every isomorphism must be a bijection, the reverse is not true. As the next example shows, not all bijections are isomorphic.

Example 5. $x \mapsto x^3$ is a bijection from $\mathbb{R}$ to $\mathbb{R}$, but it is not an isomorphism from $(\mathbb{R}, +)$ to $(\mathbb{R}, +)$, because $(x + y)^3 \neq x^3 + y^3$ in general.

3 Isomorphism Theorem

Being isomorphic is a very strong condition. If A and B are isomorphic, they are essentially the same structure. In particular, first order logic cannot distinguish between them.

Theorem 1 (Isomorphism Theorem)**.** *Let A and B be structures over the same vocabulary and $h : D^A \to D^B$ be an isomorphism. Then for all $\varphi \in FORM$ and $\alpha : VARS \to D^A$, we have $A, \alpha \models \varphi$ iff $B, (h \circ \alpha) \models \varphi$.*

Corollary 1. *Let $A \simeq B$ and let φ be a sentence. Then $A \models \varphi$ iff $B \models \varphi$.*

We will prove the theorem using structural induction on formulas. Note that $h \circ \alpha$ is the function composition of h and α, obtained by first applying α and then applying h. So it is a binding from $VARS$ to D^B that maps variable x to $h(\alpha(x))$. We first obtain the following useful result: applying the binding $(h \circ \alpha)$ to a term t is the same as first applying the binding α to t and then applying h to $\overline{\alpha}(t)$.

Lemma 1. *Let $h : D^A \to D^B$ be a homomorphism, and $\alpha : VARS \to D^A$ be an assignment. Then for all terms t, we have $h(\overline{\alpha}(t)) = \overline{h \circ \alpha}(t)$.*

Proof. By structural induction on terms.

- *Basis:* t is a variable x.

$$h(\overline{\alpha}(x)) = h(\alpha(x)) = (h \circ \alpha)(x) = \overline{h \circ \alpha}(x)$$

- *Inductive Step:* t is $f(t_1, \ldots, t_k)$.

$$
\begin{array}{rlr}
& h(\overline{\alpha}(f(t_1, \ldots, t_k))) & \\
= & h(f^A(\overline{\alpha}(t_1), \ldots, \overline{\alpha}(t_k))) & \text{(Definition of } \overline{\alpha}\text{)} \\
= & f^B(h(\overline{\alpha}(t_1)), \ldots, h(\overline{\alpha}(t_k))) & \text{(Defn. of homomorphism)} \\
= & f^B(\overline{h \circ \alpha}(t_1), \ldots, \overline{h \circ \alpha}(t_k)) & \text{(Induction Hypothesis)} \\
= & \overline{h \circ \alpha}(f(t_1, \ldots, t_k)) & \text{(Definition of } \overline{\alpha}\text{)}
\end{array}
$$

□

Proof of the Isomorphism Theorem: Let h be an isomorphism from A to B. We must show that $A, \alpha \models \varphi$ iff $B, (h \circ \alpha) \models \varphi$. We will use structural induction on formulas.

- *Basis:* φ is $P(t_1, \ldots, t_n)$.

$$
\begin{array}{rlr}
& A, \alpha \models P(t_1, \ldots, t_n) & \\
\text{iff} & (\overline{\alpha}(t_1) \ldots \overline{\alpha}(t_n)) \in P^A & \text{(Definition of } \models\text{)} \\
\text{iff} & (h(\overline{\alpha}(t_1)) \ldots h(\overline{\alpha}(t_n))) \in P^B & (h \text{ is a homomorphism}) \\
\text{iff} & (\overline{h \circ \alpha}(t_1) \ldots \overline{h \circ \alpha}(t_n)) \in P^B & \text{(By the lemma above)} \\
\text{iff} & B, (h \circ \alpha) \models P(t_1, \ldots, t_n) & \text{(Definition of } \models\text{)}
\end{array}
$$

- *Inductive Step:*
  - $\varphi = \neg\psi$.

$$
\begin{array}{rlr}
& A, \alpha \models \neg\psi & \\
\text{iff} & A, \alpha \not\models \psi & \text{(Semantics of } \neg\text{)} \\
\text{iff} & B, (h \circ \alpha) \not\models \psi & \text{(Induction Hypothesis)} \\
\text{iff} & B, (h \circ \alpha) \models \neg\psi & \text{(Semantics of } \neg\text{)}
\end{array}
$$

  - $\varphi = (\psi \wedge \theta)$.

$$
\begin{array}{rlr}
& A, \alpha \models (\psi \wedge \theta) & \\
\text{iff} & A, \alpha \models \psi \text{ and } A, \alpha \models \theta & \text{(Semantics of } \wedge\text{)} \\
\text{iff} & B, (h \circ \alpha) \models \psi \text{ and } B, (h \circ \alpha) \models \theta & \text{(Induction Hypothesis)} \\
\text{iff} & B, (h \circ \alpha) \models (\psi \wedge \theta) & \text{(Semantics of } \wedge\text{)}
\end{array}
$$

- $\varphi = (\forall x)\psi$.

$$
\begin{array}{lll}
 & A, \alpha \models (\forall x)\psi & \\
\text{iff} & \text{for all } a \in D^A,\ A, \alpha[x \mapsto a] \models \psi & \text{(Semantics of } \forall) \\
\text{iff} & \text{for all } a \in D^A,\ B, h \circ (\alpha[x \mapsto a]) \models \psi & \text{(Induction Hypothesis)} \\
\text{iff} & \text{for all } a \in D^A,\ B, (h \circ \alpha)[x \mapsto h(a)] \models \psi & (h \circ [x \mapsto a] = [x \mapsto h(a)]) \\
\text{iff} & \text{for all } b \in D^B,\ B, (h \circ \alpha)[x \mapsto b] \models \psi & (h \text{ is a bijection}) \\
\text{iff} & B, (h \circ \alpha) \models (\forall x)\psi & \text{(Semantics of } \forall)
\end{array}
$$

This concludes the proof.