Proof If $T \models \phi$, then there is n such that if G is a graph and $G \models \psi_n$, then $G \models \phi$. Thus, $p_N(\phi) \geq p_N(\psi_n)$ and by Lemma 2.4.3, $\lim_{N\to\infty} p_N(\phi) = 1$. On the other hand, if $T \not\models \phi$, then, because T is complete, $T \models \neg\phi$ and $\lim_{N\to\infty} p_N(\neg\phi) = 1$ so $\lim_{N\to\infty} p_N(\phi) = 0$.

Ehrenfeucht–Fraïssé Games

The type of back-and-forth constructions we did in Theorems 2.4.1 and 2.4.2 will appear several times in the book. It is useful to recast constructions as games. We will do this in a bit more generality. Let $\mathcal{L}$ be a language and $\mathcal{M} = (M, \ldots)$ and $\mathcal{N} = (N, \ldots)$ be two $\mathcal{L}$-structures with $M \cap N = \emptyset$. If $A \subseteq M$, $B \subseteq N$ and $f : A \to B$, we say that f is a *partial embedding* if $f \cup \{(c^{\mathcal{M}}, c^{\mathcal{N}}) : c \text{ a constant of } \mathcal{L}\}$ is a bijection preserving all relations and functions of $\mathcal{L}$.

We will define an infinite two-player game $G_\omega(\mathcal{M}, \mathcal{N})$. We will call the two players player I and player II; together they will build a partial embedding f from M to N. A play of the game will consist of ω stages. At the ith-stage, player I moves first and either plays $m_i \in M$, challenging player II to put m_i into the domain of f, or $n_i \in N$, challenging player II to put n_i into the range. If player I plays $m_i \in M$, then player II must play $n_i \in N$, whereas if player I plays $n_i \in M$, then player II must play $m_i \in M$. Player II wins the play of the game if $f = \{(m_i, n_i) : i = 1, 2, \ldots\}$ is the graph of a partial embedding.

A *strategy* for player II in $G_\omega(\mathcal{M}, \mathcal{N})$ is a function τ such that if player I's first n moves are $c_1, \ldots, c_n$, then player II's nth move will be $\tau(c_1, \ldots, c_n)$. We say that player II uses the strategy τ in the play of the game if the play looks like:

Player I	Player II
c_1	
	$\tau(c_1)$
c_2	
	$\tau(c_1, c_2)$
c_3	
	$\tau(c_1, c_2, c_3)$
$\vdots$	$\vdots$

We say that τ is a *winning strategy* for player II, if for any sequence of plays $c_1, c_2, \ldots$ player I makes, player II will win by following τ. We define strategies for player I analogously.

For example, suppose that $\mathcal{M}, \mathcal{N} \models$ DLO. Then, player II has a winning strategy. Suppose that up to stage n they have built a partial embedding $g : A \to B$. If player I plays $a \in M$, then player II plays $b \in N$ such that the cut b makes in B is the image of the cut of a in A under g. Similarly, if player I plays $b \in N$, player II plays $a \in M$ such that the cut of a in A is the image under g^{-1} of the cut of b in B. This can be done as in the proof of Theorem 2.4.1.

Proposition 2.4.5 *If $\mathcal{M}$ and $\mathcal{N}$ are countable, then the second player has a winning strategy in $G_\omega(\mathcal{M},\mathcal{N})$ if and only if $\mathcal{M} \cong \mathcal{N}$.*

Proof If $\mathcal{M}$ and $\mathcal{N}$ are isomorphic, then player II can win by playing according to the isomorphism.

Suppose that player II has a winning strategy. Let $m_0, m_1, \ldots$ list M and $n_0, n_1, \ldots$ list N. Consider a play of the game where the second player uses the winning strategy and the first player plays $m_0, n_0, m_1, n_1, m_2, n_2, \ldots$. If f is the partial embedding built during this play of the game then the domain of f is M and the range of f is N. Thus, f is an isomorphism.

By weakening the game, we can, for suitable languages, give a characterization of elementary equivalence. Fix $\mathcal{L}$ a finite language with no function symbols, and let $\mathcal{M}$ and $\mathcal{N}$ be $\mathcal{L}$-structures. We define a game $G_n(\mathcal{M},\mathcal{N})$ for $n = 1, 2, \ldots$. The game will have n rounds. On the ith round player I plays first and either plays $a_i \in M$ or $b_i \in N$. On player II's turn, if player I played $a_i \in M$, then player II must play $b_i \in N$, and if player I plays $b_i \in N$, then player II must play $a_i \in M$. The game stops after the nth round. Player II wins if $\{(a_i, b_i) : i = 1, \ldots, n\}$ is the graph of a partial embedding from $\mathcal{M}$ into $\mathcal{N}$. We call $G_n(\mathcal{M},\mathcal{N})$ an *Ehrenfeucht–Fraïssé game.*

Our goal is to prove the following theorem.

Theorem 2.4.6 *Let $\mathcal{L}$ be a finite language without function symbols and let $\mathcal{M}$ and $\mathcal{N}$ be $\mathcal{L}$-structures. Then, $\mathcal{M} \equiv \mathcal{N}$ if and only if the second player has a winning strategy in $G_n(\mathcal{M},\mathcal{N})$ for all n.*

Before proving this, we will need several lemmas.

Lemma 2.4.7 *One of the players has a winning strategy in $G_n(\mathcal{M},\mathcal{N})$.*

Proof (sketch) This follows from Zermelo's theorem that in any two-person finite length game of perfect information without ties one of the players has a winning strategy (see [10] 1.7.1). It also follows from the determinacy of closed games (see [52]). We outline the proof. Suppose that player II does not have a winning strategy. Then, there is some move player I can make in round one so that player II has no move available to force a win. Player I makes that move. Now, whatever player II does, there is still a move that if made by player I means that player II cannot force a win. Player I makes that move and continues in this way. On the last round, there is still a move possible so that player II has no winning move. Player I makes that move and wins. This informally describes a winning strategy for player I (the strategy can be summarized as "avoid losing positions").

We inductively define depth(ϕ), the *quantifier depth* of an $\mathcal{L}$-formula ϕ, as follows:

depth(ϕ) = 0 if and only if ϕ is quantifier-free;
depth($\neg\phi$) = depth(φ);
depth($\phi \wedge \psi$) = depth($\phi \vee \psi$) = max{depth(ϕ), depth(ψ)};
depth($\exists v\ \phi$) = depth(ϕ) + 1.

We say that $\mathcal{M} \equiv_n \mathcal{N}$ if $\mathcal{M} \models \phi \Leftrightarrow \mathcal{N} \models \phi$ for all sentences of depth at most n. We will show that player II has a winning strategy in $G_n(\mathcal{M}, \mathcal{N})$ if and only if $\mathcal{M} \equiv_n \mathcal{N}$. We first argue that there are only finitely many inequivalent formulas of a fixed quantifier depth.

Lemma 2.4.8 *For each n and l, there is a finite list of formulas $\phi_1, \ldots, \phi_k$ of depth at most n in free variables $x_1, \ldots, x_l$ such that every formula of depth at most n in free variables $x_1, \ldots, x_l$ is equivalent to some ϕ_i.*

Proof We first prove this for quantifier-free formulas. Because $\mathcal{L}$ is finite and has no constant symbols, there are only finitely many atomic $\mathcal{L}$-formulas in free variables $x_1, \ldots, x_l$. Let $\sigma_1, \ldots, \sigma_s$ list all such formulas.

If ϕ is a Boolean combination of formulas $\tau_1, \ldots, \tau_s$, then there is S a collection of subsets of $\{1, \ldots, s\}$ such that

$$\models \phi \leftrightarrow \bigvee_{X \in S} \Big(\bigwedge_{i \in X} \tau_i \wedge \bigwedge_{i \notin X} \neg\tau_i \Big)$$

(see Exercise 1.4.1). This gives a list of 2^{2^s} formulas such that every Boolean combination of $\tau_1, \ldots, \tau_s$ is equivalent to a formula in this list. In particular, because quantifier free formulas are Boolean combinations of atomic formulas, there is a finite list of depth-zero formulas such that every depth-zero formula is equivalent to one in the list.

Because formulas of depth $n+1$ are Boolean combinations of $\exists v\phi$ and $\forall v\phi$ where ϕ has depth at most n, the lemma follows by induction.

We can give a characterization of $\equiv_n$ using Ehrenfeucht–Fraïssé games. Theorem 2.4.6 will follow immediately.

Lemma 2.4.9 *Let $\mathcal{L}$ be a finite language without function symbols and $\mathcal{M}$ and $\mathcal{N}$ be $\mathcal{L}$-structures. The second player has a winning strategy in $G_n(\mathcal{M}, \mathcal{N})$ if and only if $\mathcal{M} \equiv_n \mathcal{N}$.*

Proof We prove this by induction on n.

Suppose that $\mathcal{M} \equiv_n \mathcal{N}$. Consider a play of the game where in round one player I plays $a \in M$. (The case where player I plays $b \in N$ is similar.) We claim that there is $b \in \mathcal{N}$ such that $\mathcal{M} \models \phi(a) \Leftrightarrow \mathcal{N} \models \phi(b)$ whenever depth(ϕ) $< n$. Let $\phi_0(v), \ldots, \phi_m(v)$ list, up to equivalence, all formulas of depth less than n. Let $X = \{i \leq m : \mathcal{M} \models \phi_i(a)\}$, and let $\Phi(v)$ be the formula

$$\bigwedge_{i \in X} \phi_i(v) \wedge \bigwedge_{i \notin X} \neg\phi_i(v).$$

Then, $\text{depth}(\exists v\ \Phi(v)) \leq n$ and $\mathcal{M} \models \Phi(a)$; thus, there is $b \in N$ such that $\mathcal{N} \models \Phi(b)$. Player II plays b in round one.

If $n = 1$, the game has now concluded and $a \mapsto b$ is a partial embedding so player II wins. Suppose that $n > 1$.

Let $\mathcal{L}^* = \mathcal{L} \cup \{c\}$, where c is a new constant symbol. View $\mathcal{M}$ and $\mathcal{N}$ as $\mathcal{L}^*$-structures $(\mathcal{M}, a)$ and $(\mathcal{N}, b)$ where we interpret the new constant as a and b, respectively. Because

$$\mathcal{M} \models \phi(a) \Leftrightarrow \mathcal{N} \models \phi(b)$$

for $\phi(v)$ an $\mathcal{L}$-formula with $\text{depth}(\phi) < n$, $(\mathcal{M}, a) \equiv_{n-1} (\mathcal{N}, b)$. By induction, player II has a winning strategy in $G_{n-1}((\mathcal{M}, a), (\mathcal{N}, b))$. If player I's second play is d, player II responds as if d was player I's first play in $G_{n-1}((\mathcal{M}, a), (\mathcal{N}, b))$ and continues playing using this strategy, that is, in round i player I has plays $a, d_2, \ldots, d_i$, then player II plays $\tau(d_2, \ldots, d_i)$, where τ is his winning strategy in $G((\mathcal{M}, a), (\mathcal{N}, b))$. Let $f : X \to N$ be the function built by this play of the game. Because τ is a winning strategy, f^* is partial $\mathcal{L}^*$-embedding. Extend f^* to $f : X \cup \{a\} \to N$ by $f(a) = b$. Because f^* preserves $\mathcal{L}$-formulas with an additional constant denoting a in $\mathcal{M}$ and b in $\mathcal{N}$, f is a partial $\mathcal{L}$-embedding. Thus a winning strategy for player II can be summarized as: given player I's first play a, find b such that $(\mathcal{M}, a) \equiv_{n-1} (\mathcal{N}, b)$ and follow the winning strategy of $G_{n-1}((\mathcal{M}, a), (\mathcal{N}, b))$.

On the other hand, suppose that $\mathcal{M} \not\equiv_n \mathcal{N}$. Because formulas of depth at most n are Boolean combinations of formulas of the form $\exists v\ \phi(v)$ where $\text{depth}(\phi) < n$, $\mathcal{M}$ and $\mathcal{N}$ must disagree about a formula of this type. Without loss of generality, we may assume that $\mathcal{M} \models \exists v\ \phi(v)$ and $\mathcal{N} \models \forall v \neg\phi(v)$ where $\text{depth}(\phi) < n$. We claim that player I has a winning strategy. In round one player I plays $a \in M$ such that $\mathcal{M} \models \phi(a)$. Suppose that player II responds with $b \in N$. Let $(\mathcal{M}, a)$ and $(\mathcal{N}, b)$ be as above. Then $(\mathcal{M}, a) \not\equiv_{n-1} (\mathcal{N}, b)$ and, by induction, player I has a winning strategy in $G_{n-1}((\mathcal{M}, a), (\mathcal{N}, b))$. Player I continues playing as if just starting a game of $G_{n-1}((\mathcal{M}, a), (\mathcal{N}, b))$. The function f^* played starting at the second move will not be a partial $\mathcal{L}^*$-embedding so the whole function played is not a partial $\mathcal{L}$-embedding.

We give one application of Theorem 2.4.6. Let $\mathcal{L} = \{<\}$. Let T be the $\mathcal{L}$-theory that asserts $<$ is a linear order and $\forall x \exists y \exists z\ (y < x < z \wedge \forall w\ (w \leq y \vee w = x \vee w \geq z))$. T is the theory of discrete orderings with no top or bottom element.

Suppose that $\mathcal{N} \models T$. For $a, b \in N$ say aEb if b is the nth successor or predecessor of a for some natural number n. Then, E is an equivalence relation. Each E-class is a linear order that looks like $(\mathbb{Z}, <)$. If aEb, $\neg(aEc)$, and $a < c$, then $b < c$. Thus, the E-classes are linearly ordered and every model of T is of the form $(L \times \mathbb{Z}, <)$, where L is a linear order and $<$ is

Ehrenfeucht-Fraïssé games

Throughout this handout we will, for simplicity, be working in a finite language without function symbols. Given two structures M and N in such a language we can detect elementary equivalence in terms of *games*.

DEFINITION 1.1. Given two models M and N and a natural number $n \in \mathbb{N}$ we define a game as follows. It is a two-player game in which two players, player I and player II, move in turn. Player I starts and the game ends after n rounds, so after both players have played n moves. A move by a player consists of picking an element from one of the two structures. Player I has complete freedom and can pick an element from whichever structures he likes, but player II always has to reply by picking an element from the other structure (that is, player II is not allowed to respond by picking an element from the same structure as the one player I just played in). So if in round i player I chooses an element $a_i \in M$, player II replies by picking an element $b_i \in N$, and if in round i player I chooses an element $b_i \in N$, then player II replies by picking an element $a_i \in M$. After n rounds the two players have constructing two sequences $\langle a_1, \dots, a_n \rangle$ and $\langle b_1, \dots, b_n \rangle$ of elements from M and N, respectively. Player II wins if $\{(a_i, b_i) \colon 1 \leq i \leq n\}$ is a well-defined injective function $f \colon \{a_1, \dots, a_n\} \to N$ and this function is moreover a local isomorphism (Marker says: partial embedding); otherwise player I wins. We denote this game $\mathcal{G}_n(M, N)$ and we call it an Ehrenfeucht-Fraïssé game.

We have:

THEOREM 1.2. *Let L be a finite language without function symbols and let M and N be L-structures. Then $M \equiv N$ if and only if the second player has a winning strategy in $\mathcal{G}_n(M, N)$ for all n.*

We give one application of this theorem. Let $L = \{<\}$. Let T be the L-theory that asserts that $<$ is a linear order and

$$\forall x \, \exists y \, \exists z \, (\, y < x < z \wedge \forall w \, (\, w \leq y \vee w = x \vee w \geq z \,)\,).$$

T is the theory of a discrete ordering with no top or bottom element.

Suppose that $N \models T$. For $a, b \in N$ say aEb if b is the nth successor or predecessor of a for some natural number n. Then E is an equivalence relation. Each E-class is a linear order that looks like $(\mathbb{Z}, <)$. If $aEb, \neg(aEc)$, and $a < c$, then $b < c$. Thus, the E-classes are linearly ordered and every model of T is of the form $(L \times \mathbb{Z}, <)$, where L is a linear order and $<$ is the lexicographic order on $L \times \mathbb{Z}$ (that is, $(a, n) < (b, n)$ if $a < b$, or both $a = b$ and $n < m$). Also, every linear order of this form is a model of T.

PROPOSITION 1.3. *The theory of discrete linear orders with no top or bottom element is a complete theory. In particular, $(\mathbb{Z}, <) \models \varphi$ if and only if $T \models \varphi$ for all L-sentences φ.*

PROOF. Let M be the ordered set of integers $(\mathbb{Z}, <)$, and let N be $L \times \mathbb{Z}$ with the lexicographic order where L is any linearly ordered set.

We claim that $M \equiv N$. We must show that player II has a winning strategy in $\mathcal{G}_n(M, N)$ for all n.

If $a, b \in \mathbb{Z}$, we define the distance between a and b to be $\text{dist}(a, b) = |b - a|$, and if $x = (i, a), y = (j, b) \in L \times \mathbb{Z}$, we define the distance to be $\text{dist}(x, y) = |b - a|$ if $i = j$ and $\text{dist}(a, b) = \infty$ if $i \neq j$. The problem for player II is that player I can play elements that are infinitely far apart in N and force player II to play elements that are finitely far apart in M. Because player II knows how long the game will last, player II can play elements sufficiently far apart to avoid conflicts. Player II will try to ensure:

> (†) After m rounds of $\mathcal{G}_n(M, N)$ we have $a_i < a_j$ iff $b_i < b_j$ and $a_i = a_j$ iff $b_i = b_j$ and $\min(\text{dist}(a_i, a_j), 2^{n-m}) = \min(\text{dist}(b_i, b_j), 2^{n-m})$.

By doing this, player II will win because after n rounds there will be a local isomorphism.

We argue that player II can always choose a move to preserve (†). In round 1, player II chooses an arbitrary element and (†) holds. Suppose that we have played m rounds and (†) holds, and the moves played so far have been $a_1, \ldots, a_m$ in M and $b_1, \ldots, b_m$ in N. Suppose that player I plays $b \in L \times \mathbb{Z}$. There are several cases to consider.

(1) $b < b_i$ for all i. Suppose b_j is the smallest element of the b_i. Then choose $a = a_j - \min(\text{dist}(b, b_j), 2^{n-m-1})$.
(2) $b_i < b < b_j$ for some i and j. Choose i and j such that $b_i < b < b_j$ and there are no b_k such that $b_i < b_k < b_j$.
   (a) If $\text{dist}(b, b_i) < 2^{n-m-1}$, then put $a = a_i + \text{dist}(b, b_i)$.
   (b) If $\text{dist}(b, b_j) < 2^{n-m-1}$, then put $a = a_j - \text{dist}(b, b_j)$.
   (c) If $\text{dist}(b, b_i) \geq 2^{n-m-1}$ and $\text{dist}(b, b_j) \geq 2^{n-m-1}$, then $\text{dist}(b_i, b_j) \geq 2^{n-m}$ and $\text{dist}(a_i, a_j) \geq 2^{n-m}$. Put $a = a_i + 2^{n-m-1}$.
(3) If $b > b_i$ for all i. Suppose b_j is the biggest element of the b_i. Then choose $a = a_j + \min(\text{dist}(b, b_j), 2^{m-n-1})$.

This explains the strategy if player I plays $b \in L \times \mathbb{Z}$. The case where player I plays $a \in \mathbb{Z}$ is analogous and left to the reader. □