Introduction to Ehrenfeucht's Game

Schoolweek OzsL
Amsterdam, December 1999

Kees Doets

September 25, 2002

Contents

Chapter 1

The Game

The notion of an Ehrenfeucht-Fraïssé game provides a simple characterization of elementary equivalence with straightforward generalizations to several languages other than first-order, which, for simple models (such as linear orderings, trees), is easy to apply. Besides, it is almost the only technique available in finite-model theory (where Compactness and Löwenheim-Skolem are of no use).

The following introduction to the subject is mainly focussed on the general theory. There are some 30 exercises, marked with ♣.

1.1 Basics

1.1 Models. For the time being, a *model* is a couple $\mathcal{A} = (A, \mathrm{R}^{\mathcal{A}})$ where A is a (usually, non-empty) set and $\mathrm{R}^{\mathcal{A}} \subseteq A \times A$ is a binary relation on A.

Examples we'll often come across:

- $\omega = (\mathbb{N}, <)$,
- $\zeta = (\mathbb{Z}, <)$,
- $\eta = (\mathbb{Q}, <)$,
- $\lambda = (\mathbb{R}, <)$.
- $\mathrm{L}_n = (\{0, \ldots, n-1\}, <)$.

1.2 Isomorphism and Local Isomorphism. An *isomorphism* between models $\mathcal{A} = (A, R)$ and $\mathcal{B} = (B, S)$ is a bijection $h : A \to B$ such that for all $a, a' \in A$: $aRa' \Leftrightarrow h(a)Sh(a')$.

A *local* or *partial isomorphism* between $\mathcal{A}$ and $\mathcal{B}$ is a finite relation h with $\mathrm{Dom}(h) \subseteq A$, $\mathrm{Ran}(h) \subseteq B$, that preserves equality and relation; i.e., such that for all $a, a' \in A$ and $b, b' \in B$ with $(a, b), (a', b') \in h$:

- $a = a' \Leftrightarrow b = b'$,
- $aRa' \Leftrightarrow bSb'$.

(Equivalently: h is a finite injection h with $\mathrm{Dom}(h) \subseteq A$, $\mathrm{Ran}(h) \subseteq B$, and for all $a, a' \in \mathrm{Dom}(h)$: $aRa' \Leftrightarrow h(a)Sh(a')$.)

1.3 Examples.

1. The empty relation $\emptyset$ is a local isomorphism between every two models.
2. Every (finite) part of a (local) isomorphism is a local isomorphism.
3. A composition of local isomorphisms is a local isomorphism. I.e.: if $g : A \to B$ is a local isomorphism between $\mathcal{A}$ and $\mathcal{B}$, and $h : B \to C$ is a local isomorphism between $\mathcal{B}$ and $\mathcal{C}$, then $h \circ g$ (the map $a \mapsto h(g(a))$, where $a \in \mathrm{Dom}(g)$ and $g(a) \in \mathrm{Dom}(h)$) is a local isomorphism between $\mathcal{A}$ and $\mathcal{C}$.
4. The relation $\{(0,0),(2,e),(5,\pi)\}$ is a local isomorphism between $(\mathbb{Z},<)$ and $(\mathbb{R},<)$.

The last example illustrates that a local isomorphism doesn't need to be part of an isomorphism.

1.4 Lemma. *A local isomorphism is the same as an isomorphism between submodels.*

I.e.: h is a local isomorphism between (A,R) and (B,S) iff it is an isomorphism between $(\mathrm{Dom}(h), R|\mathrm{Dom}(h))$ and $(\mathrm{Ran}(h), S|\mathrm{Ran}(h))$.

1.5 Ehrenfeucht game. Any two models $\mathcal{A}$, $\mathcal{B}$, together with an integer $n \in \mathbb{N}$, determine an *Ehrenfeucht Game* of *length* n.

It is played by two players: Spoiler and Duplicator. In a play of the game, Spoiler and Duplicator move alternatingly, until n moves have been made by each player.

One pair of moves consists of first Spoiler choosing an element from one of the models, and next Duplicator choosing an element from the other model.

At the end of such a play, the n pairs of moves build a finite relation between A and B. Duplicator *wins* iff this finite relation happens to be a local isomorphism. (In the opposite case, Spoiler wins; every play of the game is won by one of the players: a draw is not possible.)

Remarks.

1. Repeating previous moves is not excluded —and even necessary if there are few elements. But to repeat moves unnecessarily is not a smart thing to do for Spoiler, and if Spoiler doesn't repeat a move, it is best for Duplicator not to repeat moves either.
2. It may become evident that Spoiler wins before all $2n$ moves have been played. But for Duplicator to win, all $2n$ moves have to be executed.

3. The intuition behind the game is, that Duplicator aims at showing that the models in some way look alike; it is Spoiler's goal to spot differences. The longer the game, the easier it can be for Spoiler and the harder for Duplicator to win.

4. Duplicator wins a game of length 0 immediately: there are no moves, the relation built up is empty, and the empty relation is a local isomorphism between any two models.

5. In some cases below, the game is played in a situation with an empty model. In that case we agree that a player who cannot move loses. Thus, if $\mathcal{A}$ is empty and $\mathcal{B}$ isn't, then Spoiler wins a game of positive length by playing an element of B. However, if $A = B = \emptyset$, Duplicator wins.

Example. The length 3 game on $\zeta := (\mathbb{Z}, <)$ and $\lambda := (\mathbb{R}, <)$. Suppose Spoiler and Duplicator play as follows:

	S	D	S	D	S	D
$\mathbb{Z}$		2	0			5
$\mathbb{R}$	e			0	π	

The end result is $\{(0,0),(2,e),(5,\pi)\}$, which happens to be a local isomorphism. Duplicator has won.

The obvious question in this example is: did Duplicator won by luck, or is he clever?

1.6 Winning Strategies. A *strategy* for a player is a rule that tells him how to play in every position of the game in which he has to move.

For instance, a strategy for Spoiler in the length n-game on $\mathcal{A}$ and $\mathcal{B}$ is a function that assigns, to every relation $\{(a_1,b_1),\ldots,(a_m,b_m)\} \subseteq A \times B$ for which $0 \leqslant m < n$, an element in $A \cup B$.

A strategy σ for Spoiler is *winning* if Spoiler wins every play in which he uses σ, no matter what Duplicator does.

The notion of a winning strategy for Duplicator is defined analogously.

Example, continued. So the question is: has Duplicator a winning strategy in the length-3 game on ζ and λ?

Try to answer this, and the same question for the games of lengths 2 and 4 on these structures.

A Few Games. An intuition for the game can be developed by just playing it. Figure out whom of the players has a winning strategy in the length-3 games on the following models, and try to describe it.

1. ω and η;

   ω and ζ;

   ω and $\omega + \omega$ (the ordered sum of two copies of ω).

2. L_6 and L_7;
   L_7 and L_8.
3. $\omega + L_1 + \omega^\star$ (where * inverts the ordering) and $\omega + L_2 + \omega^\star$;
   $\omega + L_2 + \omega^\star$ and $\omega + L_3 + \omega^\star$.
4. L_7 and $\omega + \omega^\star$.
5. ω and $\omega + \zeta$.

1.2 Elementary Properties

1.7 Notation. $\mathsf{D}(\mathcal{A}, \mathcal{B}, n)$ expresses that Duplicator has a winning strategy in the length-n game on $\mathcal{A}$ and $\mathcal{B}$.

One glance a Theorem 2.4 (p. 12) explains that this is the notion we'll be interested in.

1.8 Lemma.

1. $\mathsf{D}(\mathcal{A}, \mathcal{B}, n) \wedge m \leqslant n \;\Rightarrow\; \mathsf{D}(\mathcal{A}, \mathcal{B}, m)$,
2. $\mathsf{D}(\mathcal{A}, \mathcal{B}, n) \;\Rightarrow\; \mathsf{D}(\mathcal{B}, \mathcal{A}, n)$,
3. $\mathcal{A} \cong \mathcal{B} \;\Rightarrow\; \forall n\, \mathsf{D}(\mathcal{A}, \mathcal{B}, n)$,
4. $\mathsf{D}(\mathcal{A}, \mathcal{B}, n) \wedge \mathsf{D}(\mathcal{B}, \mathcal{C}, n) \;\Rightarrow\; \mathsf{D}(\mathcal{A}, \mathcal{C}, n)$.

1.9 Determinacy. *In every Ehrenfeucht game exactly one of the players has a winning strategy.*

Proof. Of course, both players can't have winning strategies for the same game.

A game in which one of the players has a winning strategy is called *determined*. We prove the stronger result that every 2-person game that has no draws and in which all plays are finite is determined. (Thus, it is neither necessary that all plays have the *same* finite length, nor that players have to move alternatingly.)

Suppose given such a game with players I and II.

Let T be the set of all *positions* that can occur while playing it. In particular, there is an *initial* position $t_0 \in T$. For $t \in T$, there are 3 possibilities: it is I's turn to move at t, it is II's, or t is *terminal*: game over, in which case the rules of the game determine which of the two players has won.

For $s, t \in T$, write $s \twoheadrightarrow t$ if the position t can be reached from s by one move of the player whose turn it is to move in s. Thus, a play of the game has the form

$$t_0 \twoheadrightarrow t_1 \twoheadrightarrow \cdots \twoheadrightarrow t_n$$

where t_n is terminal.

Let us call a position $t \in T$ *determined* if one of the players has a winning strategy for the subgame that starts at t. Trivially, terminal positions are determined (the winning strategy of the player who has won is doing nothing.) The theorem states that the initial position t_0 is determined.

Claim. *If $s \in T$ is not determined, then some $t \twoheadleftarrow s$ is not determined.*

From this, the theorem follows: if the initial position t_0 isn't determined, there is some $t_1 \twoheadleftarrow t_0$ that isn't, hence some $t_2 \twoheadleftarrow t_1$ is not determined, etc etc; and we end up with an infinite sequence $t_0 \twoheadrightarrow t_1 \twoheadrightarrow t_2 \twoheadrightarrow \cdots$, contradicting the finiteness assumption on plays.

To prove the Claim, assume $s \in T$ is not determined. As hypothesis for a proof by contradiction, suppose that every $t \twoheadleftarrow s$ is determined.

First, note that s can't be terminal, since terminal positions are trivially determined. Thus, one of the players has to move in position s; and we may as well assume that this is player I.

(a) There exists $t \twoheadleftarrow s$ such that I has a winning strategy $\sigma = \sigma_t$ in t. Then I has a winning strategy in s as well: it consists of executing the move $s \twoheadrightarrow t$, followed by the winning strategy σ_t. Contradiction.

(b) There doesn't exist $t \twoheadleftarrow s$ such that I has a winning strategy in t. Then, by assumption (every $t \twoheadleftarrow s$ is determined), II must have a winning strategy τ_t in every position $t \twoheadleftarrow s$. It follows that II has a winning strategy in s: it consists of first waiting what I's move $s \twoheadrightarrow t$ will be (s is not terminal and it's I's turn to move), followed by the winning strategy τ_t. Contradiction again. ⊣

1.10 Corollary. (Zermelo, Euwe) *In the game of chess, either White has a winning strategy, or Black has a strategy with which he cannot lose.*

This application clearly shows the purely theoretical nature of the determinacy proof.

1♣ Winning strategies for Duplicator and transfer. The existence of a winning strategy for Duplicator can be used to transfer truth from one model to the other. (The explanation of this phenomenon follows in Chapter 2.)

1. Assume that $\mathsf{D}((A,R),(B,S),2)$, and that the relation R is symmetric ($\forall a_1, a_2 \in A(a_1 R a_2 \Rightarrow a_2 R a_1)$. Then S is symmetric as well.

2. Assume that $\mathsf{D}((A,R),(B,S),3)$, and that the relation R is dense. (R is *dense* if $\forall a, b \in A(aRb \Rightarrow \exists c \in A(aRc \wedge cRb)$.) Then S is dense too.

3. Assume that $\mathsf{D}((A,R),(B,S),3)$, and that R is confluent. (R is *confluent* if $\forall a, b_1, b_2 \in A(aRb_1 \wedge aRb_2 \Rightarrow \exists c \in A(b_1 Rc \wedge b_2 Rc))$.) Then also S is confluent.

Conversely, every specific first-order sentence that holds in $\mathcal{A}$ and is false in $\mathcal{B}$ can be transformed into a winning strategy for Spoiler for an Ehrenfeucht game of suitable length. (Of course, this follows from the above statement using determinacy. But the direct argument is illuminating.)

2♣

1. Suppose that R is symmetric, but S isn't. Describe a winning strategy for Spoiler in the length-2 game on $\mathcal{A} = (A,R)$ and $\mathcal{B} = (B,S)$.

2. Suppose that R is dense but S isn't. Describe a winning strategy for Spoiler in the length-3 game.

3. What about the case for confluency?

1.11 Proposition *Suppose that $\mathcal{A}$ has n elements.*

1. *If $\mathsf{D}(\mathcal{A}, \mathcal{B}, n)$, then there exists an embedding of $\mathcal{A}$ into $\mathcal{B}$.*

2. *If $\mathsf{D}(\mathcal{A}, \mathcal{B}, n+1)$, then $\mathcal{A} \cong \mathcal{B}$.*

Proof. 1: Let Spoiler enumerate the elements of A and Duplicator use his winning strategy. 2: An isomorphism is the same as a surjective embedding. ⊣

3♣ Find a simple condition on n and the number of elements of A and B that is both necessary and sufficient in order that $\mathsf{D}((A, \emptyset), (B, \emptyset), n)$ (both relations empty) holds.

1.3 A Few Examples

In the case that $\mathcal{A}$ and $\mathcal{B}$ are linear orderings, a statement that $\mathsf{D}(\mathcal{A}, \mathcal{B}, n)$ can often be shown using induction on n.

1.12 Notation. If $<$ linearly orders A, and $a \in A$, then the notation $a\uparrow$ is used for the submodel of $(A, <)$ with universe $\{x \in A \mid a < x\}$.

Similarly, $a\downarrow$ is the submodel of $(A, <)$ that has universe $\{x \in A \mid x < a\}$.

E.g., $-1\uparrow$ may denote the submodel $(\mathbb{N}, <)$ of $(\mathbb{Z}, <)$.

Remarks.

- Thus, if $a \in A$, then $(A, <) = a\downarrow + \{a\} + a\uparrow$.
- If a is the greatest (resp., the least) element of A, then $a\uparrow$ ($a\downarrow$) is *empty*.

1.13 Splitting Lemma. *For linear orderings $\mathcal{A}$ and $\mathcal{B}$, $\mathsf{D}(\mathcal{A}, \mathcal{B}, n+1)$ holds iff, both*

("forth") $\forall a \in A\, \exists b \in B\, [\mathsf{D}(a\downarrow, b\downarrow, n) \wedge \mathsf{D}(a\uparrow, b\uparrow, n)]$, *and*

("back") $\forall b \in B\, \exists a \in A\, [\mathsf{D}(a\downarrow, b\downarrow, n) \wedge \mathsf{D}(a\uparrow, b\uparrow, n)]$.

1.14 Example. For every n, we have that $\mathsf{D}(\lambda, \eta, n)$.
(In this particular case, Duplicator's winning strategy doesn't depend on n.)

1.15 Proposition. $k, m \geqslant 2^n - 1 \Rightarrow \mathsf{D}(\mathrm{L}_k, \mathrm{L}_m, n)$.

Proof. Induction w.r.t. n, using Lemma 1.13.
Basis. $n = 0$.
Trivial, since the empty relation always is a local isomorphism, and this is the relation built up after 0 moves. (If you find this tricky, just check the case for $n = 1$.)
Induction step.
Induction hypothesis: the statement holds for n.
Now suppose that $k, m \geqslant 2^{n+1} - 1$. In order that $\mathsf{D}(\mathrm{L}_k, \mathrm{L}_m, n+1)$, it suffices (according to Lemma 1.13) to show that ("forth") for every element i of L_k there is an element j in L_m such that $\mathsf{D}(i\downarrow, j\downarrow, n)$ and $\mathsf{D}(i\uparrow, j\uparrow, n)$, and conversely ("back": for every j in L_m there should be an i in L_k such that $\mathsf{D}(i\downarrow, j\downarrow, n)$ and $\mathsf{D}(i\uparrow, j\uparrow, n)$ — but the situation is symmetric in k and m so it suffices to only check the "forth"-claim).

Thus, suppose that i is an element of L_k. Distinguish 3 cases, depending on whether i is located "left", "right", or "in the middle" of L_k.
(i) $i\downarrow$ has $< 2^n - 1$ elements.
Pick j in L_m such, that $i\downarrow \cong j\downarrow$. (This is possible, since $m \geqslant 2^{n+1} - 1$.) Then $\mathsf{D}(i\downarrow, j\downarrow, n)$. Furthermore (since $k, m \geqslant 2^{n+1} = (2^n - 1) + 1 + (2^n - 1)$), $i\uparrow$ and $j\uparrow$ have at least $2^n - 1$ (in fact, at least 2^n) elements, and so $\mathsf{D}(i\uparrow, j\uparrow, n)$ follows from induction hypothesis.
(ii) $i\uparrow$ has $< 2^n - 1$ elements.
Pick j such that $i\uparrow \cong j\uparrow$, and argue as under (i).
(iii) $i\downarrow$ and $i\uparrow$ both have $\geqslant 2^n - 1$ elements.

Claim. *$j \in \mathrm{L}_m$ exists such that both $j\downarrow$ and $j\uparrow$ have $\geqslant 2^n - 1$ elements.*

Proof. Because of $m \geqslant 2^{n+1}$, and $2^{n+1} - 1 = (2^n - 1) + 1 + (2^n - 1)$. ⊣

Pick such a j. By induction hypothesis, $\mathsf{D}(i\downarrow, j\downarrow, n)$ and $\mathsf{D}(i\uparrow, j\uparrow, n)$ hold. ⊣

1.16 Proposition. *If $m \geqslant 2^n - 1$, then $\mathsf{D}(\omega + \omega^\star, \mathrm{L}_m, n)$; more generally, for every linear ordering α: $\mathsf{D}(\omega + (\zeta \cdot \alpha) + \omega^\star, \mathrm{L}_m, n)$.*

Proof. Another induction. ⊣

1.17 Lemma.

1. $\mathsf{D}(\alpha_1, \beta_1, n) \wedge \mathsf{D}(\alpha_2, \beta_2, n) \;\Rightarrow\; \mathsf{D}(\alpha_1 + \alpha_2, \beta_1 + \beta_2, n)$.
2. *More generally: if I is a linearly ordered set and for all $i \in I$ α_i and β_i are orderings s.t. $\mathsf{D}(\alpha_i, \beta_i, n)$, then $\mathsf{D}(\sum_{i\in I} \alpha_i, \sum_{i \in I} \beta_i, n)$.*

1.18 Proposition. *For all $n \in \mathbb{N}$:*

1. $\mathsf{D}(\omega, \omega + \zeta, n)$;

   more generally: for any α, $\mathsf{D}(\omega, \omega + \zeta \cdot \alpha, n)$, *2.* $\mathsf{D}(\zeta, \zeta+\zeta, n)$*;*

more generally: for any α*:* $\mathsf{D}(\zeta, \zeta\cdot\alpha, n)$.

Proof. 1: $\omega = (2^n-1)+\omega$, $\omega+\zeta = (\omega+\omega^\star)+\omega$. Use Proposition 1.16 and Lemma 1.17.1.
2: Similar. ⊣

B_m is the (unordered, rooted) binary tree all of whose branches have length m. This tree can be represented as the set of finite sequences of 0's and 1's of length $< m$, partially ordered by $s \prec t \equiv s$ is an initial segment of t. The length-0 sequence is the root in this tree.

The following is reminiscent of Proposition 1.15 (p. 6), but its proof is somewhat harder.

1.19 Proposition. $m, k \geqslant 2^n - 1 \ \Rightarrow\ \mathsf{D}(B_m, B_k, n)$.

Proof. Induction w.r.t. n. The case $n=0$ (or $n=1$) is trivial. For the induction step, suppose that $m, k \geqslant 2^{n+1}-1$. The reader is urged to draw pictures.

Let $a \in B_m$ be the first move of Spoiler.

In the linear ordering-case, an element induces a splitting of the ordering in (that element and) *two* halves, and we can use Lemma 1.13. In the present tree-case, the element a can be used to split B_m in *three* (or four) parts:

- the element a,
- the two *top-subtrees*, the roots of which are the two immediate successors of a (these trees are empty if a happens to be maximal),
- the poset $a\swarrow = \{t \mid a \not\preceq t\}$ that consists of the linear ordering $a\!\downarrow = \{t \mid t \prec a\}$ plus the "side-trees" sprouting from $a\!\downarrow$ ($a\swarrow$ being empty if a happens to be the root of B_m).

  Notation: for $i \prec a$, $T_i = \{t \mid i \prec t \wedge (t \not\preceq a \wedge a \not\preceq t)\}$ denotes the side-tree from $a\swarrow$, the root of which is the immediate successor of i that is $\not\preceq a$.

As in the proof of 1.15, distinguish the following cases.

(i) $a\!\downarrow$ has $\leqslant 2^n-1$ elements.

Duplicator chooses $b \in B_k$ such that $b\!\downarrow \cong a\!\downarrow$.

It now suffices to indicate that Duplicator wins the n-round games on corresponding parts in the decompositions of the two trees B_m and B_k that are induced by a and b.

Since $|a\!\downarrow| = |b\!\downarrow| \leqslant 2^n-1$, the top-trees above a and b have height $\geqslant 2^n-1$; thus, by IH, Duplicator has winning strategies for the n-round games on the two pairs of top-trees.

On the posets $a\swarrow$, $b\swarrow$, Duplicator counters in $a\!\downarrow$ and $b\!\downarrow$ using the isomorphism between these linear orderings, and he counters in side-trees that correspond under this isomorphism using winning strategies for the n-round games. These strategies exist according to IH (note that all side-trees have height $\geqslant 2^n$).

(ii) The subtree with root a has height $\leqslant 2^n$.

It clearly suffices to decompose the two trees as $B_m = T_1 \cup (B_m - T_1)$ resp., $B_k = T_2 \cup (B_k - T_2)$, in such a way that $a \in T_1$, $\mathsf{D}(T_1, T_2, n+1)$, and $\mathsf{D}(B_m - T_1, B_k - T_2, n)$.

Choose $\overline{a} \preceq a$ such that the subtree T_1 with root $\overline{a}$ has height *exactly* 2^n. Decompose B_m in T_1 and $\overline{a} \swarrow = B_m - T_1$, the latter consisting of the linear ordering $\overline{a}\downarrow$ of length $\geqslant 2^n - 1$ and the side-trees T_i $(i \prec \overline{a})$, all of them of height $\geqslant 2^n$.

Similarly decompose B_k into some subtree T_2 of height *exactly* 2^n with root $\overline{b}$ and the rest $\overline{b} \swarrow = B_k - T_2$, which consists of, again, the linear ordering $\overline{b}\downarrow$ of length $\geqslant 2^n - 1$ with side-trees T_j $(j \prec \overline{b})$ of height $\geqslant 2^n$.

We now have that $T_1 \cong T_2$, and hence, $\mathsf{D}(T_1, T_2, n+1)$.

Now $\mathsf{D}(\overline{a} \swarrow, \overline{b} \swarrow, n)$ follows from Proposition 1.15 and IH, as follows. Fix:

- a winning strategy σ for Duplicator for the n-round game between $\overline{a}\downarrow$ and $\overline{b}\downarrow$ (using 1.15),
- a winning strategy σ_{ij} (for each $i \prec \overline{a}$, $j \prec \overline{b}$) for Duplicator for the n-round game between the side-trees T_i and T_j (using IH).

Moves by Spoiler in $\overline{a}\downarrow$, $\overline{b}\downarrow$ are now countered by Duplicator using σ. A move i^+ of Spoiler in, say, T_i, $i \prec \overline{a}$, is countered as follows. First, σ produces an answer $j \prec \overline{b}$ to i (and possibly earlier moves or elements considered in $\overline{a}\downarrow$, $\overline{b}\downarrow$). Next, σ_{ij} produces an answer j^+ to i^+ (and possibly earlier moves in T_i, T_j).

(iii) $a\downarrow$ has $\geqslant 2^n - 1$ elements and the subtree with root a has height $\geqslant 2^n$. Duplicator selects an element $b \in B_k$ with the same properties.

Decompose B_m and B_k as under (i) using a resp., b. It suffices to see that Duplicator has winning strategies for the n-round games on corresponding parts in the decompositions.

Use IH for the top-trees above a and b (which have height $\geqslant 2^n - 1$). To see that $\mathsf{D}(a \swarrow, b \swarrow, n)$, again, fix:

- a winning strategy σ for Duplicator for the n-round game between $a\downarrow$ and $b\downarrow$ (using 1.15),
- for every $i \prec a$, $j \prec b$, a winning strategy σ_{ij} for Duplicator for the n-round game between the side-trees T_i and T_j (using IH).

The strategy followed by Duplicator is the same as under (ii). ⊣

The (finite) binary tree $\mathcal{C} = \mathcal{C}(\mathrm{L}_m, B_{l_0}, \ldots, B_{l_m})$ is the disjoint union of the linear ordering $\mathrm{L}_m = (\{0, \ldots, m-1\}, <)$ and the $m+1$ finite binary trees $B_{l_0}, \ldots, B_{l_m}$ by letting each $i \in \mathrm{L}_m$ $(i = 0, \ldots, m-2)$ be the immediate predecessor of the root of B_{l_i} and, moreover, by letting the greatest element $m - 1 \in \mathrm{L}_m$ be the immediate predecessor of the roots of both $B_{l_{m-1}}$ and B_{l_m}.

Similarly, the (infinite) binary tree $\mathcal{D} = \mathcal{D}(\omega, B_{j_0}, B_{j_1}, B_{j_2}, \ldots)$ is the disjoint union of the linear ordering ω and $B_{j_0}, B_{j_1}, B_{j_2}, \ldots$ by letting each $i \in \omega$ be the immediate predecessor of the root of B_{j_i}.

1.20 Proposition. $m, l_0, \ldots, l_m, j_0, j_1, j_2, \ldots \geqslant 2^n - 1 \;\Rightarrow\; \mathsf{D}(\mathcal{C}, \mathcal{D}, n)$.

Proof. Induction w.r.t. n.

For the induction step, assume that $m, l_0, \ldots, l_m, j_0, j_1, j_2, \ldots \geqslant 2^{n+1} - 1$.

In the following we show that ("forth"), for every $a \in \mathcal{C}$, we can decompose $\mathcal{C} = \mathcal{C}_1 \cup \mathcal{C}_2$ and $\mathcal{D} = \mathcal{D}_1 \cup \mathcal{D}_2$ such that $a \in \mathcal{C}_1$, $\mathsf{D}(\mathcal{C}_1, \mathcal{D}_1, n+1)$, and $\mathsf{D}(\mathcal{C}_2, \mathcal{D}_2, n)$, and ("back") a similar claim holds for every $b \in \mathcal{D}$. Draw pictures illustrating these decompositons!

If $a \in \mathcal{C}$ is the first move of Spoiler, we may, by swapping $B_{l_{m-1}}$ and B_{l_m}, wlog assume that $a \in \mathcal{C}_1 = \mathcal{C} - B_{l_m} = \mathrm{L}_m \cup B_{l_0} \cup \cdots B_{l_{m-1}}$. $\mathcal{C}$ is decomposed into $\mathcal{C}_1$ and $\mathcal{C}_2 = B_{l_m}$.

Choose $k \geqslant 2^{n+1} - 1$ in ω so large that, if $b \in \mathcal{D}$ happened to be the first move of Spoiler, then $k \not\leq b$. $\mathcal{D}$ is decomposed into into $\mathcal{D}_1 \cup \mathcal{D}_2$, where $\mathcal{D}_1 = k \swarrow$ is the initial segment $\mathrm{L}_k = \{0, \ldots, k-1\}$ of ω plus the side-trees $B_{j_0}, \ldots, B_{j_{k-1}}$ (thus, k has been chosen such that a possibly first move of Spoiler in $\mathcal{D}$ happened in $\mathcal{D}_1$), and $\mathcal{D}_2$ is the rest: the subtree with root k, which consists of the part $\{k, k+1, k+2, \ldots\}$ of ω, plus side-trees $B_{j_k}, B_{j_{k+1}}, B_{j_{k+2}}, \ldots$

We now have that $\mathsf{D}(\mathcal{C}_1, \mathcal{D}_1, n+1)$, as in parts (ii) and (iii) of the proof of Proposition 1.19, using the result of this proposition.

What we need, furthermore, is that $\mathsf{D}(B_{l_m}, \mathcal{D}_2, n)$. To see this, simply rewrite B_{l_m}, in the notation for $\mathcal{C}$ that is explained immediately above 1.20, as $B_{l_m} = \mathcal{C}(\mathrm{L}_{2^n-1}, B_{l_0}, \ldots, B_{l_{2^n-1}})$, where $l_0, \ldots, l_{2^n-1} \geqslant 2^n$. From this, it is clear that $\mathsf{D}(B_{l_m}, \mathcal{D}_2, n)$ follows by IH. ⊣

4♣ $(\mathcal{P}(A), \subseteq)$ is the model in which $\mathcal{P}(A)$ is the set of all subsets of A. Show:

1. If A and B are infinite, then for all $n \in \mathbb{N}$: $\mathsf{D}((\mathcal{P}(A), \subseteq), (\mathcal{P}(B), \subseteq), n)$.
2. If $|A|, |B| \geqslant 2^n$, then $\mathsf{D}((\mathcal{P}(A), \subseteq), (\mathcal{P}(B), \subseteq), n)$.

Chapter 2

Logic

2.1 Main Theorem

Models, Formulas. There is no reason to stick to just one relation; models are allowed to have the form $\mathcal{A} = (A, R, S, \dots)$ with *finitely many* relations (of any arity) $R, S, \dots$ over the universe A. Consequently, atomic formulas have the form: $x{=}y$, $\mathrm{R}(x_1, \dots, x_n)$, $\mathrm{S}(x_1, \dots, x_m)$,... From these, (first-order) formulas are built using connectives and quantifiers.

2.1 Quantifier Rank. The *quantifier rank* $\mathrm{qr}(\varphi)$ of a (first-order) formula φ is the maximum number of nested quantifiers in φ. I.e.:

1. for atomic φ, $\mathrm{qr}(\varphi) = 0$,
2. $\mathrm{qr}(\neg\varphi) = \mathrm{qr}(\varphi)$,
3. $\mathrm{qr}(\varphi \to \psi) = \mathrm{qr}(\varphi \wedge \psi) = \mathrm{qr}(\varphi \vee \psi) = \mathrm{qr}(\varphi \leftrightarrow \psi) = \max(\mathrm{qr}(\varphi), \mathrm{qr}(\psi))$,
4. $\mathrm{qr}(\forall x\varphi) = \mathrm{qr}(\exists x\varphi) = \mathrm{qr}(\varphi) + 1$.

Examples. The quantifier rank of $\forall x(\exists y x\mathrm{R}y \wedge \exists y \neg x\mathrm{R}y)$ is 2. The confluency sentence (Exercise 1 p. 5) $\forall x, y_1, y_2(x\mathrm{R}y_1 \wedge x\mathrm{R}y_2 \to \exists z(y_1\mathrm{R}z \wedge y_2\mathrm{R}z))$ has rank 4. Its logical equivalent $\forall y_1, y_2(\exists x(x\mathrm{R}y_1 \wedge x\mathrm{R}y_2) \to \exists z(y_1\mathrm{R}z \wedge y_2\mathrm{R}z))$ has rank 3.

2.2 Equivalence and n-Equivalence. Models $\mathcal{A}$ and $\mathcal{B}$ are (elementary, first-order) *equivalent* if they have the same true first-order sentences. Notation: $\mathcal{A} \equiv \mathcal{B}$.

They are *n-equivalent* if they have the same true sentences of rank $\leqslant n$. Notation: $\mathcal{A} \equiv^n \mathcal{B}$.

Thus: $\mathcal{A} \equiv \mathcal{B}$ holds iff, for all $n \in \mathbb{N}$, we have that $\mathcal{A} \equiv^n \mathcal{B}$.

Example. L_2 and L_3 are not 2-equivalent: a distinguishing rank-2 sentence is $\exists y(\exists x(x < y) \wedge \exists z(y < z))$. These models *are* equivalent w.r.t. sentences with two quantifiers only.

2.3 Lemma. *For every k and n there are, up to logical equivalence, only finitely many formulas with at most $x_1, \ldots, x_k$ free and quantifier rank $\leqslant n$.*

Proof. This is due to the fact that a vocabulary consists of finitely many relation symbols. The proof uses induction w.r.t. n, using disjunctive normal forms. In the induction step for $n+1$, use that the quantifier rank $\leqslant n+1$ formulas with at most $x_1, \ldots x_k$ free are, up to logical equivalence, generated using the booleans from (i) the quantifier rank $\leqslant n$ formulas with at most $x_1, \ldots x_k$ free, and (ii) the formulas $\exists x_{k+1}\varphi$ with φ of quantifier rank $\leqslant n$ and at most $x_1, \ldots x_k, x_{k+1}$ free. (Thus, the IH is used for $k+1$ instead of k.) $\dashv$

2.4 Main Theorem. $\mathsf{D}(\mathcal{A}, \mathcal{B}, n) \Leftrightarrow \mathcal{A} \equiv^n \mathcal{B}$.

In Ehrenfeucht's game, it is clear from the proof below that the moves are "meant" as values for bound variables. But formulas are built, next to quantifiers, from connectives as well. So: what is left of the connectives in the Ehrenfeucht game?

2.5 Corollary. $\forall n\, \mathsf{D}(\mathcal{A}, \mathcal{B}, n) \Leftrightarrow \mathcal{A} \equiv \mathcal{B}$.

2.6 Examples. $(\mathbb{Q}, <) \equiv (\mathbb{R}, <)$ (By Example 1.14 p. 6, for all n, $\mathsf{D}(\eta, \lambda, n)$); $(\mathbb{N}, <) \equiv (\mathbb{N} + \mathbb{Z}, <)$ (By Proposition 1.18.1 p. 7, for all n, $\mathsf{D}(\omega, \omega + \zeta, n)$).

Proof. (Of 2.4.) The following generalisation can be handled using induction. N.B.: as in the proof of Lemma 2.3, k needs to be kept variable: in the induction step for k and $n+1$, IH is applied for n and $k+1$.

2.7 Lemma. *For all n and every finite relation $h = \{(a_1, b_1), \ldots, (a_k, b_k)\} \subseteq A \times B$, the following conditions are equivalent:*

1(n) Duplicator *has a winning strategy in position h in the length-$(k+n)$ game on the models with n more rounds to go,*

2(n) *for all formulas $\varphi = \varphi(x_1, \ldots, x_k)$ with $\mathrm{qr}(\varphi) \leqslant n$, we have that*

$$\mathcal{A} \models \varphi[a_1, \ldots, a_k] \text{ iff } \mathcal{B} \models \varphi[b_1, \ldots, b_k].$$

Basis: $n = 0$.
In fact, the following are equivalent:

1(0) h is a local isomorphism (= Duplicator has a win with 0 rounds to go),

2′(0) for all *atomic* formulas φ with $\mathrm{Var}(\varphi) \subseteq \{x_1, \ldots, x_k\}$ we have that

$$\mathcal{A} \models \varphi[a_1, \ldots, a_k] \text{ iff } \mathcal{B} \models \varphi[b_1, \ldots, b_k],$$

2(0) for all formulas φ with $\mathrm{Var}(\varphi) \subseteq \{x_1, \ldots, x_k\}$ and $\mathrm{qr}(\varphi) = 0$ we have that

$$\mathcal{A} \models \varphi[a_1, \ldots, a_k] \text{ iff } \mathcal{B} \models \varphi[b_1, \ldots, b_k].$$

N.B.: For $h = \emptyset$ all three conditions hold: 1(0), since $\emptyset$ is a local isomorphism between every two models, and $2'(0)$ and 2(0) since there are no atomic or rank-0 formulas without free variables.

Induction step.
Induction hypothesis: the equivalence $1(n) \Leftrightarrow 2(n)$ holds.
$1(n+1) \Rightarrow 2(n+1)$. Suppose that Duplicator has a win in position h in the length-$(k+n+1)$ game with $n+1$ more rounds to go. Here follows proof that, for all $\varphi = \varphi(x_1, \dots, x_k)$ with $\mathrm{qr}(\varphi) \leqslant n+1$, $\mathcal{A} \models \varphi[a_1, \dots, a_k]$ iff $\mathcal{B} \models \varphi[b_1, \dots, b_k]$.

Induction w.r.t. φ.

Since h is also a win for Duplicator in the length-$(k+n)$ game, by induction hypothesis we get the required equivalence for formulas of rank $\leqslant n$, in particular, for atomic ones. Induction steps for the connectives are effortless. Quantifier case:

Assume $\mathcal{A} \models \exists x_{k+1}\psi[a_1, \dots, a_k]$, where $\mathrm{qr}(\psi) \leqslant n$. Thus, let $a_{k+1} \in A$ be such that $\mathcal{A} \models \psi[a_1, \dots, a_k, a_{k+1}]$.

Consider a_{k+1} as a move of Spoiler in position h in the length-$(k+n+1)$ game. Since h is a win for Duplicator, there is a move $b_{k+1} \in B$ bringing Duplicator to a position $h' := h \cup \{(a_{k+1}, b_{k+1})\}$ that, again, is a win for Duplicator.

By induction hypothesis applied to h', we have that for all formulas $\varphi = \varphi(x_1, \dots, x_k, x_{k+1})$ with $\mathrm{qr}(\varphi) \leqslant n$: $\mathcal{A} \models \varphi[a_1, \dots, a_k, a_{k+1}]$ is true iff $\mathcal{B} \models \varphi[b_1, \dots, b_k, b_{k+1}]$. In particular, $\mathcal{B} \models \psi[b_1, \dots, b_k, b_{k+1}]$, and hence $\mathcal{B} \models \exists x_{k+1}\psi[b_1, \dots, b_k]$.

$2(n+1) \Rightarrow 1(n+1)$. Suppose Spoiler chooses $a_{k+1} \in A$. We show that Duplicator has an answer $b_{k+1} \in B$ bringing him into a position $h' := h \cup \{(a_{k+1}, b_{k+1})\}$ in the length-$(k+n+1)$ game that is won for him. (The conclusion being, that Duplicator had a win already in position h.)

Consider the set

$$\Phi := \{\varphi(x_1, \dots, x_k, x_{k+1}) \mid \mathrm{qr}(\varphi) \leqslant n \ \wedge\ \mathcal{A} \models \varphi[a_1, \dots, a_k, a_{k+1}]\}.$$

By Lemma 2.3 there is a finite subset $\Phi' \subseteq \Phi$ such that every element of Φ has an equivalent in Φ'.

Clearly, we have that $\mathcal{A} \models \exists x_{k+1} \bigwedge \Phi'[a_1, \dots, a_k]$. By condition 2 and since $\mathrm{qr}(\exists x_{k+1} \bigwedge \Phi') \leqslant n+1$, it follows that $\mathcal{B} \models \exists x_{k+1} \bigwedge \Phi'[b_1, \dots, b_k]$. Say, $\mathcal{B} \models \bigwedge \Phi'[b_1, \dots, b_k, b_{k+1}]$.

Claim. *If* $\mathrm{qr}(\varphi) \leqslant n$*, then:* $\mathcal{A} \models \varphi[a_1, \dots, a_{k+1}]$ *iff* $\mathcal{B} \models \varphi[b_1, \dots, b_{k+1}]$.

This follows from the choice of Φ' and b_{k+1}. The required conclusion follows using the induction hypothesis. ⊣

Constants. The Ehrenfeucht game for models with constants is played exactly as before, but now, a finite relation h is called a *local isomorphism* between $(\mathcal{A}, a_1, \dots, a_k)$ and $(\mathcal{B}, b_1, \dots, b_k)$ —where $\mathcal{A}$ and $\mathcal{B}$ are purely relational— if $h \cup \{(a_1, b_1), \dots, (a_k, b_k)\}$ is a local isomorphism between $\mathcal{A}$ and $\mathcal{B}$ in the old sense. Equivalently (as above, $2'(0)$): a local isomorphism is a correspondence that preserves satisfaction of atomic formulas.

2.8 Corollary. *The Main Theorem 2.4 is valid for languages with finitely many constant symbols.*

Proof. Immediate from Lemma 2.7. Suppose that the models $\mathcal{A}' = (\mathcal{A}, a_1, \ldots, a_k)$ and $\mathcal{B}' = (\mathcal{B}, b_1, \ldots, b_k)$ expand the purely relational models $\mathcal{A}$ and $\mathcal{B}$ with k constants each. Put $h = \{(a_1, b_1), \ldots, (a_k, b_k)\}$. Then the following conditions are equivalent:

1. Duplicator has a winning strategy in the length-n game on the models $\mathcal{A}'$ and $\mathcal{B}'$,

2. Duplicator has a winning strategy in position h in the length-$(k+n)$ game on the models $\mathcal{A}$ and $\mathcal{B}$ with n more rounds to go,

3. for all formulas $\varphi = \varphi(x_1, \ldots, x_k)$ with $\mathrm{qr}(\varphi) \leqslant n$, we have that
$$\mathcal{A} \models \varphi[a_1, \ldots, a_k] \text{ iff } \mathcal{B} \models \varphi[b_1, \ldots, b_k],$$

4. for all sentences $\varphi = \varphi(\mathrm{c}_1, \ldots, \mathrm{c}_k)$ with $\mathrm{qr}(\varphi) \leqslant n$, we have that
$$\mathcal{A}' \models \varphi \text{ iff } \mathcal{B}' \models \varphi.$$

$\dashv$

Remark. You can now see what finiteness of the vocabulary is good for. For instance, let $\mathcal{B}$ be a proper elementary extension of $(\mathbb{N}, 0, 1, 2, \ldots)$. (Every proper extension of this model happens to be an elementary one.) Spoiler can already win the length 1 game on these models by choosing an element of $\mathcal{B}$ outside $\mathbb{N}$. A similar example with $(\mathbb{N}, 0, S)$ (where $S(n) := n + 1$) illustrates why you have to exclude function symbols.

Thus, from now on we can allow finitely many constant symbols.

5♣ Show: if $\mathcal{A}$ is finite and $\mathcal{A} \equiv \mathcal{B}$, then $\mathcal{A} \cong \mathcal{B}$.

6♣ Show: every two dense linear orderings without endpoints are equivalent.

7♣ Suppose that the linear ordering α can be embedded into the linear ordering β. Show that $\zeta \cdot \alpha$ can be elementarily embedded in $\zeta \cdot \beta$. (An embedding is *elementary* if it preserves all formulas.)

8♣ Cf. Lemma 1.16 (p. 7).

1. Produce, for every $n \in \mathbb{N}$, a sentence φ_n of rank n that is true of a linear ordering iff it has at least $2^n - 1$ elements.

2. Give a simple condition on m and n that is both necessary and sufficient in order that $\mathsf{D}(\omega + \omega^{\star}, \mathrm{L}_m, n)$.

Solution. 1. For a formula φ and a variable x not in φ, $\varphi^{<x}$ is the formula obtained from φ by replacing quantifiers $\forall y \cdots$ and $\exists y \cdots$ by $\forall y < x \cdots$ ($= \forall y(y < x \rightarrow \cdots)$), resp., $\exists y < x \cdots$ ($= \exists y(y < x \wedge \cdots)$).
$\varphi^{>x}$ idem.

Define $\varphi_1 = \exists x_1(x_1{=}x_1)$, $\varphi_{n+1} = \exists x_{n+1}(\varphi_n^{<x_{n+1}} \wedge \varphi_n^{>x_{n+1}})$.

2. $\mathsf{D}(\omega + \omega^\star, \mathrm{L}_m, n) \;\Leftrightarrow\; m \geqslant 2^n - 1$.

9♣ Cf. Lemma 1.15.

1. Construct, for $n \geqslant 2$ and $k < 2^n - 1$, a sentence $\psi_{n,k}$ of rank $\leqslant n$ that is true of a linear ordering iff it has exactly k elements.

2. Give a simple condition on k, m and n that is both necessary and sufficient in order that $\mathsf{D}(\mathrm{L}_k, \mathrm{L}_m, n)$.

Hint for 1. Start with $n = 2$ (then $2^2 - 1 = 3$), and $k = 1$, $k = 2$.
next, suppose $\psi_{n,k}$ defined for $n \geqslant 2$ and $k < 2^n - 1$. To construct $\psi_{n+1,k}$, distinguish $1 \leqslant k < 2^n - 1$, $k = 2^n - 1$, $2^n - 1 < k < 2^{n+1} - 2$, and $k = 2^{n+1} - 2$.

2.2 Applications

2.2.1 Definability

2.9 Definability. A (first-order) formula $\varphi = \varphi(x)$ in one free variable x (first-order-) *defines* the set $\varphi^{\mathcal{A}} = \{a \in A \mid \mathcal{A} \models \varphi[a]\}$ in $\mathcal{A}$; a formula $\psi = \psi(x, y)$ in two free variables x, y *defines* the relation $\psi^{\mathcal{A}} = \{(a, b) \in A \times A \mid \mathcal{A} \models \psi[a, b]\}$ in $\mathcal{A}$.

For every $n \in \mathbb{N}$ one can write a formula $\pi_n = \pi_n(x)$ in the language of $(\mathbb{N}, <)$ that expresses that (the value of) x has exactly n predecessors. Thus, $(\mathbb{N}, <) \models \pi_n[m]$ is true iff $m = n$. Consequently, if $A \subseteq \mathbb{N}$ is finite, it can be defined in $(\mathbb{N}, <)$ by the disjunction $\bigvee_{n \in A} \pi_n$; its complement $\mathbb{N} - A$ is defined by the negation of this formula.

A set $X \subseteq \mathbb{N}$ is called *co-finite* if $\mathbb{N} - X$ is finite. Thus: all finite and co-finite sets $\subseteq \mathbb{N}$ are definable in $(\mathbb{N}, <)$.

2.10 Proposition. *Every set definable in* $(\mathbb{N}, <)$ *is either finite or co-finite.*

Proof. Suppose that $\varphi(x)$ defines a set that is neither finite nor co-finite. Thus, $\omega = (\mathbb{N}, <) \models \forall x \exists y(x < y \wedge \varphi(y)) \wedge \forall x \exists y(x < y \wedge \neg\varphi(y))$. But, $\omega \equiv \omega + \zeta$; hence this sentence is true in the latter model as well. Therefore, some element a in the ζ-part satisfies φ, and some element b in the ζ-part satisfies $\neg\varphi$. Apply the automorphism of $\omega + \zeta$ that moves a to b; a contradiction results. $\dashv$

10♣ Suppose that X is definable in $(\mathbb{N} + \mathbb{Z}, <)$. Show that $\mathbb{Z} \subseteq X$ or $\mathbb{Z} \cap X = \emptyset$. Show that X is finite or co-finite.

2.11 Model-transformations and Translations. Suppose that $\delta = \delta(x, y)$ is a formula with x and y free.

1. For $\mathcal{A} = (A, R)$, the model $\mathcal{A}^\delta$ is defined as $\mathcal{A}^\delta = (A, \delta^\mathcal{A})$, where $\delta^\mathcal{A}$ is the relation defined by δ in $\mathcal{A}$.
2. The formula φ^δ is obtained from φ by replacing atomic subformulas $u\mathrm{R}v$ by $\delta(u, v)$. (Possibly renaming bound u and v in δ to avoid clashes.)

The following equivalence holds:

$$\mathcal{A}^\delta \models \varphi[a_1, \dots, a_n] \;\Leftrightarrow\; \mathcal{A} \models \varphi^\delta[a_1, \dots, a_n].$$

(Induction w.r.t. φ.)

E.g., we may so use the formula $\mathbf{suc}(x, y)$ (**suc** for (immediate) **suc**cessor):

$$x < y \;\wedge\; \neg\exists z(x < z \;\wedge\; z < y)$$

and its symmetric version $\mathbf{nb}(x, y) = \mathbf{suc}(x, y) \vee \mathbf{suc}(y, x)$ (**nb** for **n**eigh**b**our).

For instance, the relation of the model $(\mathbb{N}, <)^{\mathbf{suc}}$ is the successor relation defined by $n + 1 = m$.

Note:

1. If $\mathrm{qr}(\delta) = k$, then $\mathcal{A} \equiv^{n+k} \mathcal{B} \;\Rightarrow\; \mathcal{A}^\delta \equiv^n \mathcal{B}^\delta$; and, hence:
2. $\mathcal{A} \equiv \mathcal{B} \;\Rightarrow\; \mathcal{A}^\delta \equiv \mathcal{B}^\delta$.

11♣ Consider the "circle"-model $\mathrm{C}_m = (\{0, \dots, m-1\}, R)$, where R is defined by $iRj :\equiv i + 1 = j \;\vee\; (i = m - 1 \;\wedge\; j = 0)$. (Visualize by drawing points $0, \dots, m-1$ on a circle.) $\rho(x, y) = \mathbf{suc}(x, y) \;\vee\; (\neg\exists z(z < y) \;\wedge\; \neg\exists z(x < z))$.

1. Check that $\mathrm{C}_m = \mathrm{L}_m^\rho$ and $\zeta^{\mathbf{suc}} = (\omega + \omega^\star)^\rho$.
2. Give a sufficient condition in order that $\mathsf{D}(\mathrm{C}_m, \mathrm{C}_k, n)$.
3. Idem, for $\mathsf{D}(\mathrm{C}_m, \zeta^{\mathbf{suc}}, n)$.

2.12 Proposition. *The ordering $<$ of $\mathbb{N}$ is not definable in $\omega^{\mathbf{suc}}$.*

Proof. Suppose $\varphi(x, y)$ defines the ordering in $\omega^{\mathbf{suc}}$. Then the sentence $\Phi = \forall x\forall y(x \neq y \to (\varphi(x, y) \leftrightarrow \neg\varphi(y, x))$ holds in $\omega^{\mathbf{suc}}$. But, $\omega \equiv \omega + \zeta + \zeta$; hence $\omega^{\mathbf{suc}} \equiv (\omega + \zeta + \zeta)^{\mathbf{suc}}$; and so the sentence holds in the latter model as well. However, $(\omega + \zeta + \zeta)^{\mathbf{suc}}$ has an automorphism that interchanges the two ζ-copies. Picking (a value for) x in one and (a value for) y in the other results in a contradiction. ⊣

Note that the ordering $<$ of $\mathbb{N}$ is the *transitive closure* of the successor relation of $\omega^{\mathbf{suc}}$. Thus: transitive closures are not first-order definable. In fact, this is true already on finite models:

2.12' Proposition. *There is no uniform first-order definition of the ordering on the models $\mathrm{L}_m^{\mathbf{suc}}$.*

Proof. Suppose that φ constitutes such a definition. Then the above sentence Φ holds in every model $\mathrm{L}_m^{\mathbf{suc}}$, and hence, $\Phi^{\mathbf{suc}}$ holds in every L_m. It follows that $\omega + \zeta + \zeta + \omega^\star \models \Phi^{\mathbf{suc}}$, and hence, $(\omega + \zeta + \zeta + \omega^\star)^{\mathbf{suc}} \models \Phi$. A contradiction arises as in the earlier proof above. ⊣

A sentence *defines* the class of models in which it is true. Relative to this notion of definability:

2.13 Proposition. *There is no first-order definition of finiteness for linear orderings.*

Proof. By Lemma 1.16 (p. 7). ⊣

More interestingly:

2.14 Proposition. *There is no first-order definition of finiteness for binary trees.*

Proof. By the result of Proposition 1.20 (p. 9). ⊣

Note that 2.13 also has an easy proof using Compactness, but 2.14 hasn't since the class of binary trees isn't elementary. (An infinite model of the theory of all binary trees B_m need not be a tree at all; e.g., it might very well be non-wellfounded.) The Ehrenfeucht game technique is essential for the proof.

Connectivity. Suppose that $a, b \in A$, where (A, R) is some model (graph).

A *path connecting* a *with* b is a finite sequence $a_1 = a, \ldots, a_n = b$ s.t. for all i, $1 \leqslant i < n$: $a_i R a_{i+1}$.

The model is *connected* if for all $a, b \in A$, there is a path connecting a with b.

2.15 Proposition. *There is no first-order definition of connectivity. This is true even on the class of finite (graph) models.*

Proof. For arbitrary models, this can be proved using compactness. An infinite example is $\zeta^{\mathbf{suc}} \equiv (\zeta + \zeta)^{\mathbf{suc}}$ (since $\zeta \equiv \zeta + \zeta$); the first model is connected whereas the second one is not.

On finite models, one needs game theory. Here is an ingenious proof based on the fact that a (rank 2) first-order formula $\rho^2 = \rho^2(x, y)$ exists (a modification of the formula ρ from Exercise 11), satisfying

$$\mathrm{L}_n \models \rho^2[i, j] \;\Leftrightarrow\; i + 2 = j \;\vee\; (i = n - 2 \;\wedge\; j = 0) \;\vee\; (i = n - 1 \;\wedge\; j = 1).$$

$\rho(x, y)$ says:

"$x < y$ and there is exactly one element between them,

or: x is greatest element and there is exactly one element $< y$,

or: there is exactly one lement $> x$ and y is least element".

Picture some cases $n = 4, 5, \ldots$ and note: $\mathrm{L}_n^{\rho^2}$ is connected iff n is odd. In fact, for n odd: $\mathrm{L}_n^{\rho^2} \cong \mathrm{C}_n$; and for $n = 2m$ even: $\mathrm{L}_n^{\rho^2} \cong \mathrm{C}_m + \mathrm{C}_m$ (the disjoint union of two copies of C_m).

Now suppose that a quantifier rank p sentence defines connectivity. For $n = 2^{p+2} - 1$, we have that $\mathrm{L}_n \equiv^{p+2} \mathrm{L}_{n+1}$. Hence, $\mathrm{L}_n^{\rho^2} \equiv^p \mathrm{L}_{n+1}^{\rho^2}$. A contradiction follows. ⊣

2.2.2 Axiomatizability

2.16 Axiomatisation.

1. A set Σ of sentences *axiomatizes* (the theory of) a model $\mathcal{A}$ if for all sentences φ: $\mathcal{A} \models \varphi$ iff $\Sigma \models \varphi$.

   Equivalently: $\mathcal{A}$ is a model of Σ, and every sentence true of $\mathcal{A}$ follows from Σ.

2. Σ *axiomatizes* (the theory of) a class K of models if for all φ: φ is true in every model from K iff $\Sigma \models \varphi$.
   (Thus, an axiomatization for $\mathcal{A}$ is the same as one for $\{\mathcal{A}\}$.)

Remarks. If Σ defines K, then it also axiomatizes K. However, the converse doesn't hold.
Example: FLO is the class of finite linear orderings; ELO consists of the sentences expressing the properties of linear orderings, existence of endpoints, and the statements that every non-least (-last) element has an immediate predecessor (successor).

Claim. ELO *axiomatizes* FLO.

Proof. Obviously, every finite linear ordering satisfies ELO. Thus, if ELO$\models \varphi$, then φ is true of every finite linear ordering. Conversely: assume that not ELO$\models \varphi$. Then some $\mathcal{A} \models$ELO exists of which φ is false. From the definition of ELO, it is not hard to see that $\mathcal{A}$ must be a linear ordering that is either finite or has order type $\omega + \zeta \cdot \alpha + \omega^\star$ for some α. Let $n = \mathrm{qr}(\varphi)$. Then, by Lemma 1.16, $\mathcal{A} \equiv^n \mathrm{L}_{2^n-1}$. Thus, φ is false of the model L_{2^n-1} of FLO. $\dashv$

Thus, ELO axiomatizes FLO but it doesn't define it.

Using the above basic results on orderings, it is not hard to find axiomatisations for η, $\omega + \omega^\star$, ω, and ζ.

Example: ω is finitely axiomatized by the sentences stating: the properties of linear orderings, existence of a least element, every element has an immediate successor, and every non-least element has an immediate predecessor.

Proof. Obviously, ω satisfies these principles; thus every logical consequence of them is true of ω. Conversely, suppose that φ doesn't follow logically from these principles. Then a model $\mathcal{A}$ of them doesn't satisfy φ. It is not hard to see that $\mathcal{A}$ must be a linear ordering of some type $\omega + \zeta \cdot \alpha$. But, $\omega + \zeta \cdot \alpha \equiv \omega$. Thus, φ is false of ω. $\dashv$

12♣ Show: there is no finite axiomatisation for $\omega + \omega^\star$.

13♣ Assume that, among the models of Σ, there are arbitrarily big finite linear orderings. Show that $\omega + \omega^\star$ is a model of Σ.

14♣ Show: if the lineair orderings α and β are finitely axiomatizable, then so are $\alpha^\star$, $1+\alpha$, $\alpha+1$ and $\alpha+1+\beta$. However, $\alpha+\beta$ isn't necessarily finitely axiomatizable.

Successor relations. SUC consists of the following sentences: $\forall x \exists y(x\mathrm{R}y \wedge \forall z(x\mathrm{R}z \to z{=}y))$, $\forall x\exists y(y\mathrm{R}x \wedge \forall z(z\mathrm{R}x \to z{=}y))$, and (1) $\neg\exists x_1(x_1\mathrm{R}x_1)$, (2) $\neg\exists x_1\exists x_2(x_1\mathrm{R}x_2 \wedge x_2\mathrm{R}x_1)$, (3) $\neg\exists x_1\exists x_2\exists x_3(x_1\mathrm{R}x_2 \wedge x_2\mathrm{R}x_3 \wedge x_3\mathrm{R}x_1)$, ...

15♣ Show: every model for SUC is of the form $(\zeta\cdot\alpha)^{\mathbf{suc}}$.

16♣ Show: every sentence true of $\zeta^{\mathbf{suc}}$ has a finite model. In particular, $\zeta^{\mathbf{suc}}$ isn't finitely axiomatizable.

2.2.3 Partition Arguments

A class of finite models (suitably coded as sequences of symbols) is in NP if membership in the class is **N**on-deterministically Turing machine decidable in **P**olynomial time. The following result explains the relationship with second-order definability.

2.17 Theorem. (Fagin 1974) *On the class of finite models:* Σ^1_1 =NP.

This is probably the first genuine result in the field of *descriptive complexity* which has been quite succesful in relating computational complexity with logical definability, and that really got started some ten years ago. (Cf. the books by Ebbinghaus/Flum and Immerman.)

Any Σ^1_1-property of finite models whose complement is not Σ^1_1 (a candidate being the NP-complete graph property *3-colorability*) would give you that co-NP≠NP. From this, you may guess that showing something to be not Σ^1_1 is a tough nut. Restricting to *monadic*-Σ^1_1, where the relations quantified over are sets, can be more tractable. For instance, we've seen (Theorem 2.15 p. 17) that connectivity on finite models is not first-order; but in fact, something stronger holds:

2.18 Theorem. *Connectivity is not monadic-*Σ^1_1 *on the class of finite graphs.*

Proof. Suppose that $\exists X_1\cdots\exists X_n\sigma$ defines connectivity, where $\sigma=\sigma(\mathrm{R},X_1,\dots,X_n)$ is first-order and has rank n.

Let M be a finite set of models that picks an element from each n-equivalence class.

By the finite version of Ramsey's Theorem, there exists an m so large that every partition $h:[m]^2\to M$ has a homogeneous set of 2^{n+1} elements (in fact, 3 elements suffices for the argument below).

Choose $X_1,\dots,X_k$ such that $\mathcal{A}=(\mathrm{L}^{\mathbf{suc}}_m,X_1,\dots,X_k)\models\sigma$.

For $0\leqslant i<j<m$, $h(\{i,j\})\in M$ is the model that is $\equiv^n$ $[i,j)$; here and below, the interval notation $[i,j)$ is used for the corresponding submodel.

Let $Q\subseteq\{0,\dots,m-1\}$ be an 2^{n+1}-element set homogeneous for h. Say, for $i<j$ in Q, $h(\{i,j\})=\alpha$.

We now have, using a more or less self-explanatory notation (for successor-structures α and β, $\alpha+\beta$ is their disjoint union where, moreover, maxα is connected to minβ —insofar as these elements exist):

$$\begin{array}{rcl}\mathcal{A} & = & (\leftarrow,\min Q)+[\min Q,\max Q)+[\max Q,\rightarrow) \\ & \equiv^n & (\leftarrow,\min Q)+\alpha\cdot(2^{n+1}-1)^{\mathbf{suc}}+[\max Q,\rightarrow) \\ & \equiv^n & (\leftarrow,\min Q)+\alpha\cdot(\omega+\omega^\star)^{\mathbf{suc}}+[\max Q,\rightarrow) \\ & \equiv^n & (\leftarrow,\min Q)+\alpha\cdot(\omega+\zeta+\omega^\star)^{\mathbf{suc}}+[\max Q,\rightarrow) \\ & = & (\leftarrow,\min Q)+\alpha\cdot(\omega+\omega^\star)^{\mathbf{suc}}+[\max Q,\rightarrow)+\alpha\cdot\zeta^{\mathbf{suc}} \\ & \equiv^n & \mathcal{A}+\alpha\cdot\zeta^{\mathbf{suc}} \\ & \equiv^n & \mathcal{A}+\alpha\cdot C_{2^{n+2}-1}.\end{array}$$

2nd line: Q has 2^{n+1} elements.
Last line: $\mathrm{L}_{2^{n+2}-1}\equiv^{n+2}\omega+\omega^\star$; thus $\mathrm{L}^{\mathbf{suc}}_{2^{n+2}-1}\equiv^{n+1}(\omega+\omega^\star)^{\mathbf{suc}}$; connecting the (rank-1 definable) endpoints, we obtain $\mathrm{C}_{2^{n+2}-1}\equiv^n\zeta^{\mathbf{suc}}$.

That $\mathcal{A}\equiv^n\mathcal{A}+\alpha\cdot\mathrm{C}_{2^{n+3}-1}$ is a contradiction, since $\mathcal{A}$ is connected, but $\mathcal{A}+\alpha\cdot\mathrm{C}_{2^{n+3}-1}$ (the disjoint union of the line model $\mathcal{A}$ and a circle model) isn't.

Note that the model described on the third line of the above calculation is already disconnected; however, it is infinite too, and the purpose of the rest of the calculation is to produce a finite n-equivalent. ⊣

The notion of *connectivity* is monadic-Π^1_1: G is connected iff for all $U\subseteq G$: if $U\neq\emptyset$ and U is closed under the relation of G, then $U=G$. And if we allow an existential quantification over a *binary* relation, a definition can be concocted: G is connected if it has a linear ordering $<$ with the property that every non-least element y is connected with some $x<y$. Thus, connectivity is Σ^1_1.

There is some subtlety involved here. Consider the closely related notion of *reachability*: in a graph (A,R), b is *reachable* from a if there is a path connecting a with b. It turns out that (for finite models) *undirected reachability* (the notion for *undirected* graphs, that is: models (A,R) where R is symmetric; *edges* identified with pairs $(x,y),(y,x)\in R$) is simpler than the general (directed) notion. *Undirected* reachability is monadic-Σ^1_1 (Kanellakis 1986):

Proposition. *In a finite. undirected graph, b is reachable from a iff for some $X\subseteq A$: $a,b\in X$, a and b both have exactly* one *edge connecting them with (an element of) X, and every other $c\in X$ has exactly* two *such edges.*

Proof. $\Rightarrow$: Suppose that $a=a_0,\dots,a_n=b$ is a shortest path connecting a with b. Then $X=\{a_0,\dots,a_n\}$ satisfies the conditions stated.
$\Leftarrow$: If X satisfies these conditions, follow the path starting at a, using edges connecting elements of X. This path can't loop, and so it must end somewhere; the only possible endpoint being b. ⊣

However, Ajtai and Fagin showed in 1990 (using Ehrenfeucht's game coupled with probabilistic arguments) that *directed* reachability is not monadically Σ^1_1.

This suggests looking at the closure of monadic Σ^1_1 and Π^1_1 under first-order quantification.

According to Proposition 2.14 (p. 17), *finiteness* is not a first-order property of binary trees. By König's Lemma, a finitely branching tree is finite iff all of its branches are finite. Thus, finiteness is monadic-Π^1_1 on the class of finitely branching trees. However:

2.19 Theorem. *Finiteness is not monadic-Σ^1_1 on the class of binary trees.*

Finally, here is another example of a partition argument.

2.20 Theorem. *Every monadic-Σ^1_1-sentence $\sigma = \sigma(<)$ with a well-ordered model has a well-ordered model of type $< \omega^\omega$.*

Proof. Suppose that $\alpha = (A, <, X_1, \ldots, X_k)$ is a well-ordered model. It suffices to show that for every n, α has a well-ordered n-equivalent of type $< \omega^\omega$. In the following, we can forget about the sets $X_1, \ldots, X_k$ since they won't spoil the argument.

Fix n. By the Downward Löwenheim-Skolem Theorem, there is no loss of generality in assuming that A is countable. Apply induction with respect to the order type of α.

If α has only one element, then α itself is the required n-equivalent. (For, $1 < \omega^\omega$.)

Next, suppose that $\alpha = \beta + 1$. Then by induction hypothesis, β has such an n equivalent β', and $\beta' + 1 \equiv^n \beta + 1 = \alpha$ is the required equivalent. (Note that if $\beta < \omega^\omega$, then $\beta + 1 < \omega^\omega$.)

Finally, assume that α has a limit type. Let $a_0 \in \alpha$ be the least element of α. Since α is countable, there is a countable sequence $a_0 < a_1 < a_2 < \cdots$ that is unbounded in α. For $i < j$, let $h(i,j)$ be the set of rank-n sentences true in the submodel $[a_i, a_j)$. We may think of h as taking finitely many values only. By the infinite version of Ramsey's Theorem there exist $k_0 < k_1 < k_2 < \cdots$ such that all $h(k_i, k_j)$ are the same. By induction hypothesis, there is a well-ordering $\gamma < \omega^\omega$ that is an n-equivalent of every $[a_{k_i}, a_{k_j})$. Again by induction hypothesis, let β be a well-ordering of type $< \omega^\omega$ that is n-equivalent with $[a_0, a_{k_0})$. Then (by Lemma 1.17.2) $\beta + \gamma \cdot \omega \equiv^n [a_0, a_{k_0}) + \sum_i [a_{k_i}, a_{k_{i+1}}) = \alpha$, hence $\beta + \gamma \cdot \omega$ is the required n-equivalent of α. (Note that if $\beta, \gamma < \omega^\omega$, then $\beta + \gamma \cdot \omega < \omega^\omega$.) ⊣

Let Ω be the well-ordering of *all* ordinals.

2.21 Corollary. $\Omega \equiv \omega^\omega$.

Proof. Show that $\Omega \equiv^n \omega^\omega$ by induction on n. Use Lemma 1.13 and the fact that final segments of Ω (resp., ω^ω) have type Ω (resp., ω^ω). ⊣

17♣ Show that every monadic-Σ^1_1 sentence true of ω is also true of $\omega + \zeta$. Nevertheless: produce a set X of natural numbers such that no expansion of $\omega + \zeta$ is elementarily equivalent to (ω, X).

18♣ Is every monadic-Σ^1_1 sentence true of λ true of η as well?

19♣ Show: if $\alpha < \beta \leq \omega^\omega$, then $\alpha \not\equiv \beta$.

20♣ A linear ordering is *scattered* if it does not embed η. Let Σ be the least set of order types such that (i) $0, 1 \in \Sigma$, (ii) $\alpha, \beta \in \Sigma \Rightarrow \alpha + \beta \in \Sigma$, (iii) $\alpha \in \Sigma \Rightarrow \alpha \cdot \omega, \alpha \cdot \omega^\star \in \Sigma$. Show: every ordering in Σ is scattered, and every sentence with a scattered model has a model in Σ.

Hint. Use the technique of the proof of Proposition 2.20. Suppose that a certain first-order sentence of quantifier rank q is true in the scattered model $(A, <)$. Without loss of generality, assume that A is countable. Identify every submodel of $(A, <)$ with its universe. For $a, c \in A$, write $a \sim c$ in case that (i) $a < c$ and for all a', c' s.t. $a \leq a' < c' \leq c$, $(a', c') := \{b \in A \mid a' < b < c'\}$ has a q-equivalent in Σ, or (ii) $c < a$ and a similar statement holds, or (iii) $a = c$. Then $\sim$ is an equivalence. Clearly, if $a \sim c$ and $a < b < c$, then $a \sim b$. Thus, A is an ordered sum of equivalence classes $\sum_{i \in I} C_i$, where I is a certain linear ordering.

Show that the order type of I is dense.

Since $(A, <)$ is scattered, conclude that I is a singleton; i.e.: A is the only equivalence class.

Finally, show that A itself has a q-equivalent in Σ. If A has no greatest element, choose $a_0 < a_1 < a_2 < \cdots$ cofinal in A and apply Ramsey's theorem to see that $\{c \in A \mid a_0 < c\}$ has a q-equivalent in Σ. Do this also for $\{c \in A \mid c < a_0\}$, by chosing, if necessary, $b_0 = a_0 > b_1 > b_2 > \cdots$ coinitial in A.

Chapter 3

Wider Theory

3.1 Other Characterisations

3.1.1 Characteristics

For every $n \in \mathbb{N}$, the "game-theoretic behaviour" of a model $\mathcal{A}$ in length-n games can be coded into one sentence $\varepsilon^n_{\mathcal{A}}$, the *$n$-characteristic* of $\mathcal{A}$.

3.1 Characteristics. For a model $\mathcal{A}$, a finite sequence $\vec{a} = (a_1, \ldots, a_k)$ from A and an integer $n \in \mathbb{N}$, the formula $\varepsilon^n_{\vec{a}} = \varepsilon^n_{\mathcal{A},\vec{a}}(x_1, \ldots, x_k)$ with $x_1, \ldots, x_k$ free and $x_{k+1}, \ldots, x_{k+n}$ bound is defined as follows:

1. $\varepsilon^0_{\vec{a}}$ is the conjunction of all atoms and negations of atoms $\varphi = \varphi(x_1, \ldots, x_k)$ with at most $x_1, \ldots, x_k$ free such that $\mathcal{A} \models \varphi[a_1, \ldots, a_k]$;
2. $\varepsilon^{n+1}_{\vec{a}}$ is $\forall x_{k+1} \bigvee_{a \in A} \varepsilon^n_{(a_1,\ldots,a_k,a)} \wedge \bigwedge_{a \in A} \exists x_{k+1} \varepsilon^n_{(a_1,\ldots,a_k,a)}$.
3. $\varepsilon^n_{\mathcal{A}}$ is $\varepsilon^n_{\mathcal{A},\emptyset}$, where $\emptyset$ is the empty sequence (of length 0).

Remarks.
In a given *finite* vocabulary there are, for any k, only finitely many atoms in the variables $x_1, \ldots, x_k$. Thus, the formulas $\varepsilon^0_{\mathcal{A},(a_1,\ldots,a_k)}$ are genuine (finite) first-order formulas. If the number of atoms in these variables is A, there are $2A$ atoms and negations of atoms, and so there are at most 2^{2A} many formulas of the form $\varepsilon^0_{\mathcal{A},(a_1,\ldots,a_k)}$ (where $\mathcal{A}$ is any model and $(a_1, \ldots, a_k)$ is any length-k sequence from A).

The same works for $k + 1$, hence it follows that the conjunction and disjunction in forming $\varepsilon^1_{\mathcal{A},(a_1,\ldots,a_k,a)}$ are over an at most finite number of formulas. Thus, the rank-1 characteristics are first-order formulas.

Let $P_{k,n}$ be the number of n-characteristics for length-k sequences (with $x_1, \ldots, x_k$ free, in any model). Clearly, every formula $\varepsilon^1_{\mathcal{A},(a_1,\ldots,a_k)}$ can be identified with the set $\{\varepsilon^0_{\mathcal{A},(a_1,\ldots,a_k,a)} \mid a \in A\}$. It follows that $P_{k,1} \leqslant 2^{P_{k+1,0}}$.

These arguments continue throughout the hierarchy: in (the definition of) any $\varepsilon^{n+1}_{\mathcal{A},\vec{a}}$, disjunction and conjunction are over finitely many formulas, and $P_{k,n+1}$ is at most $2^{P_{k+1,n}}$.
Note furthermore that

$\varepsilon^n_{\vec{a}}$ has quantifier rank n, and
$\mathcal{A} \models \varepsilon^n_{\vec{a}}[\vec{a}]$.

3.2 Theorem. *The following conditions are equivalent:*

1. $\mathsf{D}(\mathcal{A}, \mathcal{B}, n)$,
2. $\mathcal{B} \models \varepsilon^n_{\mathcal{A}}$,
3. $\varepsilon^n_{\mathcal{B}} = \varepsilon^n_{\mathcal{A}}$.

(For the last equivalent to make sense, conjunctions and disjunctions must be considered as taken over sets: order and repetitions don't count.)

21♣ Show this.
Hint. Show using induction w.r.t. n that, more generally: a position $\{(a_1, b_1), \dots, (a_k, b_k)\}$ is a win for Duplicator , iff $\mathcal{B} \models \varepsilon^n_{\mathcal{A},(a_1,\dots,a_k)}[b_1, \dots, b_k]$, iff $\varepsilon^n_{\mathcal{B},(b_1,\dots,b_k)} = \varepsilon^n_{\mathcal{A},(a_1,\dots,a_k)}$.

22♣ Show: every quantifier-n sentence φ is logically equivalent with a finite disjunction of sentences of the form $\varepsilon^n_{\mathcal{A}}$.
In fact: $\varphi \sim \bigvee\{\varepsilon^n_{\mathcal{A}} \mid \mathcal{A} \models \varphi\}$.

3.1.2 Fraïssé

Here follows the Fraïssé-characterisation of n-equivalence (that preceded Ehrenfeucht's).

3.3 Fraïssé sequence. A *Fraïssé sequence* of length $n+1$ for $\mathcal{A}$, $\mathcal{B}$ is a sequence $I_0, \dots, I_n$ of sets of local isomorphisms between $\mathcal{A}$ and $\mathcal{B}$ such that $\emptyset \in I_n$ and for all i, $0 \leqslant i < n$: if $h \in I_{i+1}$, then

("forth") $\forall a \in A \exists b \in B\ (h \cup \{(a, b)\} \in I_i)$,

and

("back") $\forall b \in B \exists a \in A\ (h \cup \{(a, b)\} \in I_i)$.

3.4 Theorem. *The following are equivalent:*

1. $\mathsf{D}(\mathcal{A}, \mathcal{B}, n)$,
2. *there is a* Fraïssé sequence *of length* $n + 1$ *for* $\mathcal{A}$ *and* $\mathcal{B}$.

Proof. $1 \Rightarrow 2$. Assume $\mathsf{D}(\mathcal{A}, \mathcal{B}, n)$. For $0 \leqslant i \leqslant n$, let I_i be the set of positions h in which Duplicator has a winning strategy for i more rounds. (N.B.: by 1, $\emptyset \in I_n$.)

$2 \Rightarrow 1$. Duplicator takes care that, after i rounds ($0 \leqslant i \leqslant n$), a local isomorphism h has been built that is an element of I_i. ⊣

3.5 Remark. In Definition 3.3, we could equivalently require that the local isomorphisms in I_i ($0 \leqslant i \leqslant n$) consist of exactly $n - i$ ordered pairs (in particular, that $I_n = \{\emptyset\}$). This will be essential in Chapter 4, when coding such sets I as relations R by means of

$$R = \{(a_1, \ldots, a_i, b_1, \ldots, b_i) \mid \{(a_1, b_1), \ldots, (a_i, b_i)\} \in I\},$$

thereby transforming the notion of a Fraïssé sequence into the following modified version (simultaneously renumbering $i \mapsto n - i$):

2'. There is a sequence $R_0, \ldots, R_n$ of relations $R_i \subseteq A^i \times B^i$ such that R_0 is $\top$ (for "true"), and

(a) for $0 \leqslant i \leqslant n$: if $R_i(a_1, \ldots, a_i, b_1, \ldots, b_i)$, then $\{(a_1, b_1), \ldots, (a_i, b_i)\}$ is a local isomorphism,
and

(b) for $0 \leqslant i < n$: if $R_i(a_1, \ldots, a_i, b_1, \ldots, b_i)$, then

("forth") $\forall a \in A\, \exists b \in B\, R_{i+1}(a_1, \ldots, a_i, a, b_1, \ldots, b_i, b)$,
and

("back") $\forall b \in B\, \exists a \in A\, R_{i+1}(a_1, \ldots, a_i, a, b_1, \ldots, b_i, b)$.

3.1.3 Co-inductive Definability

Before giving the final characterisation, a (short) introduction into (co-) inductive definability is needed.

Suppose that Γ is a *monotone operator over I*, that is: Γ maps subsets of I to subsets of I such that

$$X \subseteq Y \subseteq I \;\Rightarrow\; \Gamma(X) \subseteq \Gamma(Y).$$

3.6 Prime Example. Our one and only example is this: I is the set of local isomorphisms between two models $\mathcal{A}$ and $\mathcal{B}$, and $\Gamma = \Gamma_{\mathcal{A},\mathcal{B}}$ is defined by

$$\Gamma(X) = \{h \mid \forall a \in A \exists b \in B (h \cup \{(a, b)\} \in X) \;\wedge\; \forall b \in B \exists a \in A (h \cup \{(a, b)\} \in X)\}.$$

Note that this operator is monotone.

3.7 Post-fixed point, Co-induction. $Y \subseteq I$ is called

1. *post-fixed point* of Γ if $Y \subseteq \Gamma(Y)$,
2. *co-inductive* if for all $X \subseteq I$: $X \subseteq \Gamma(X) \;\Rightarrow\; X \subseteq Y$.

3.8 Lemma.

1. *There is at most one co-inductive post-fixed point.*
2. *A co-inductive post-fixed point is the same as a greatest fixed point.*

Proof. 1. Trivial.

2. Suppose that Y is a co-inductive post-fixed point. Thus, $Y \subseteq \Gamma(Y)$. By monotonicity, $\Gamma(Y) \subseteq \Gamma(\Gamma(Y))$; i.e.: $\Gamma(Y)$ is a post-fixed point as well. By co-induction, $\Gamma(Y) \subseteq Y$. Thus, $\Gamma(Y) = Y$. $\dashv$

As to existence:

3.9 Theorem. *The set*

$$\Gamma{\downarrow} = \bigcup\{X \subseteq I \mid X \subseteq \Gamma(X)\}$$

is the greatest fixed point of Γ.

Proof. Suppose that X is an arbitrary post-fixed point. Then, $X \subseteq \Gamma{\downarrow}$. By monotonicity, $X \subseteq \Gamma(X) \subseteq \Gamma(\Gamma{\downarrow})$. Since X was arbitrary, it follows that $\Gamma{\downarrow} \subseteq \Gamma(\Gamma{\downarrow})$: $\Gamma{\downarrow}$ is a post-fixed point. It is co-inductive by definition. $\dashv$

3.10 Fixed point hierarchy. For all ordinals α, define $\Gamma{\downarrow}\alpha \subseteq I$ by the following recursion:

1. $\Gamma{\downarrow} 0 = I$
2. $\Gamma{\downarrow}(\alpha + 1) = \Gamma(\Gamma{\downarrow}\alpha)$
3. $\Gamma{\downarrow}\gamma = \bigcap_{\xi<\gamma} \Gamma{\downarrow}\xi$ (for limits γ).

An alternative recursion would use the single equation

$$\Gamma{\downarrow}\alpha = \bigcap_{\xi<\alpha} \Gamma(\Gamma{\downarrow}\xi)$$

where it is understood that the empty intersection denotes I.

3.11 Theorem. $\Gamma{\downarrow} = \bigcap_\alpha \Gamma{\downarrow}\alpha$.

Proof. That $\bigcap_\alpha \Gamma{\downarrow}\alpha$ is co-inductive is easy: if $X \subseteq \Gamma(X)$, it follows by induction that, for all α, $X \subseteq \Gamma{\downarrow}\alpha$. Thus, $X \subseteq \bigcap_\alpha \Gamma{\downarrow}\alpha$.

To show that $\bigcap_\alpha \Gamma{\downarrow}\alpha$ is a post-fixed point, first note that the hierarchy is decreasing: $\alpha < \beta \Rightarrow \Gamma{\downarrow}\beta \subseteq \Gamma{\downarrow}\alpha$ (induction w.r.t. β; a preliminary induction shows that $\Gamma{\downarrow}(\beta+1) \subseteq \Gamma{\downarrow}\beta$). There is an ordinal β where the hierarchy becomes stationary: $\bigcap_\alpha \Gamma{\downarrow}\alpha = \Gamma{\downarrow}\beta$ (the argument needs (i) that I is a set, and hence (ii) by the Powerset Axiom, its powerset is a set as well, and (iii) by the Substitution Axiom, the map $\alpha \mapsto \Gamma{\downarrow}\alpha$ from ordinals into this powerset cannot be injective). In particular, $\Gamma{\downarrow}\beta$ is a fixed point, and $\Gamma(\bigcap_\alpha \Gamma{\downarrow}\alpha) = \Gamma(\Gamma{\downarrow}\beta) = \Gamma{\downarrow}\beta = \bigcap_\alpha \Gamma{\downarrow}\alpha$. $\dashv$

Remark. The above treatment of greatest fixed points can be dualized for least fixed points (which are much more common). Just revert inclusions and interchange intersections and unions. A different way: the least fixed point of Γ is the greatest one of the dualized operator $X \mapsto I - \Gamma(I - X)$.

3.1.4 Fixed Point Characterisation

3.12 Theorem. *The following are equivalent:*

1. $\mathsf{D}(\mathcal{A}, \mathcal{B}, n)$,
2. $\emptyset \in \Gamma_{\mathcal{A},\mathcal{B}}\!\downarrow n$.

Proof. More generally, Duplicator has a winning strategy in position h with n more rounds to go iff $h \in \Gamma\!\downarrow n$.

Note also that if $\emptyset \in \Gamma\!\downarrow n$ (equivalently, if $\Gamma\!\downarrow n \neq \emptyset$), then $\Gamma\!\downarrow 0, \ldots, \Gamma\!\downarrow n$ is a Fraïssé sequence. ⊣

3.13 Corollary. *The following are equivalent:*

1. $\mathcal{A} \equiv \mathcal{B}$,
2. $\emptyset \in \Gamma_{\mathcal{A},\mathcal{B}}\!\downarrow \omega$.

As to the relevance of the greatest fixed point $\Gamma\!\downarrow$: see Section 3.3 (p. 29).

3.2 Variations

There are variations on the Ehrenfeucht game that are adequate with respect to languages other than first-order. For instance, to get the version for (say: monadic) second-order logic, Spoiler is allowed to also pick a *subset* of one of the models; Duplicator then is obliged to counter with a subset from the other one.

A nice variation with applications to intensional logics is the one to formulas with a bounded number of variables. (The relation with intensional logics comes from the fact that standard translations into first-order logic can be carried out with finitely many variables, depending on the type of logic considered.) From the proof of Theorem 2.4 it can be seen that the moves of the players are meant as assignments of elements to variables. Now, modify the game as follows. Let $k \in \mathbb{N}$ be a natural number. Spoiler and Duplicator are given k *pebbles* each, marked $1, \ldots, k$. A move of Spoiler now consists of placing one of her pebbles on an element of one of the two models; Duplicator counters by placing his corresponding pebble on an element of the other model. If the length of the game exceeds k, Spoiler runs out of pebbles after her k-th move. She is allowed now to re-use one of her pebbles by simply moving it to some other element (of either model). Duplicator then counters by re-using his corresponding pebble. At every stage of the play, the positions of the $2k$ pebbles determine an (at most) k-element relation between the models; and Duplicator *wins* if all of them are local isomorphisms. For the k-pebble game, there is the following

3.14 Proposition. Duplicator *has a winning strategy for the k-pebble game of length n on $\mathcal{A}$ and $\mathcal{B}$ iff $\mathcal{A}$ and $\mathcal{B}$ satisfy the same rank $\leq n$-sentences with at most k variables.*

In the context of linear orderings, 3 variables suffice.

3.15 Proposition. *If $\mathcal{A}$ and $\mathcal{B}$ are linear orderings with the same valid 3-variable sentences of rank $\leq n$, then $\mathcal{A} \equiv^n \mathcal{B}$.*

Proof. Using induction, it is shown that for every n: if g and h are the locations of at most 3×2 pebbles on $\mathcal{A}$ resp. $\mathcal{B}$ such that Duplicator has a winning strategy in the 3-pebble game of length n at position (g, h), then Duplicator has a winning strategy in the ordinary game of length n at position (g, h).
Basis: $n = 0$. Trivial.
Induction step. Assume the result for n. Suppose that Duplicator has a winning strategy in the 3-pebble game of length $n + 1$ at position (g, h). Distinguish two cases.
(i) At position (g, h), only 2×2 or less pebbles have been placed. Then each player has at least one free pebble. Thus: for every $a \in A$ there exists $b \in B$ and for every $b \in B$ there exists $a \in A$ such that Duplicator has a winning strategy in the 3-pebble game of length n at position $(g \cup \{a\}, h \cup \{b\})$. By induction hypothesis: for every $a \in A$ there exists $b \in B$ and for every $b \in B$ there exists $a \in A$ such that Duplicator has a winning strategy in the ordinary game of length n at position $(g \cup \{a\}, h \cup \{b\})$. But that means that Duplicator has a winning strategy in the ordinary game of length $n + 1$ at (h, g).
(ii) At position (g, h), all 3×2 pebbles have been used. Suppose that g consists of $a_0 < a_1 < a_2$ and h is $b_0 < b_1 < b_2$. A fortiori, Duplicator has winning strategies for the two 3-pebble games of length $n + 1$ at the two-pebble positions $((a_0, a_1), (b_0, b_1))$ and $((a_1, a_2), (b_1, b_2))$. The argument under (i) shows that Duplicator has winning strategies σ resp. τ in the ordinary games of length $n + 1$ at positions $((a_0, a_1), (b_0, b_1))$ resp. $((a_1, a_2), (b_1, b_2))$. But then, Duplicator has a winning strategy in the ordinary game of length $n + 1$ at position $((a_0, a_1, a_2), (b_0, b_1, b_2))$ as well: moves $< a_1$ or $< b_1$ are countered using σ, whereas moves $> a_1$ or $> b_1$ are countered using τ. $\dashv$

3.16 Corollary. *On the class of linear orderings, every sentence is equivalent with a three-variable sentence.*

Another modification of the game is obtained by stipulating that Duplicator wins a play in case the relation built is not a local *isomorphism* but a local *homomorphism*, which is a relation $\{(a_1, b_1), \ldots, (a_n, b_n)\} \subseteq A \times B$ such that every atomic sentence true in $(\mathcal{A}, a_1, \ldots, a_n)$ is satisfied by $(\mathcal{B}, b_1, \ldots, b_n)$ as well (but not necessarily conversely). Every local homomorphism is a function (if $a_i = a_j$, then we must also have that $b_i = b_j$), but it is not necessarily an injective one.

The resulting *homomorphism-game* relates to *positive* formulas, which are generated from the atomic ones using the logical symbols $\wedge$, $\vee$, $\forall$ and $\exists$ only (thus, $\neg$, $\rightarrow$ and $\leftrightarrow$ are not allowed).

Theorem 2.4 now modifies to the following, the proof of which can be obtained by straightforward adaptation of the former one.

3.17 Theorem. Duplicator *has a winning strategy for the length-n homomorphism game iff $\mathcal{B}$ satisfies every positive quantifier rank $\leq n$ sentence true in $\mathcal{A}$.*

For another variaton in this vein, cf. the proof of Theorem 4.4 (p. 39).

Finally, you can mix requirements. Assume that $L' = L \cup \{R\}$, where R is some n-ary relation symbol. Stipulate that Duplicator wins iff the end-product of the play is a local isomorphism with respect to L-structure, and a local homomorphism with respect to R. This determines the R-*positive game*. The game is related to so-called R-*positive* sentences, which only use $\wedge$, $\vee$, $\neg$, $\forall$ and $\exists$ and in which R occurs in the scope of an *even* number of negation symbols. (The restriction that $\rightarrow$ and $\leftrightarrow$ do not occur is needed to keep the counting of negations straight: $\rightarrow$ and $\leftrightarrow$ contain "hidden" negations.)

These variations on the basic Ehrenfeucht game have their own characterisations in terms of characteristics, Fraïssé sequences and fixed point hierarchies. E.g., as to the homomorphism game, characteristics $\pi^n_{\mathcal{A},\vec{a}}$ for $n > 0$ are built as before, but now $\pi^0_{\mathcal{A},\vec{a}}$ is the conjunction of all (negationless) atoms satisfied by $\vec{a}$ in $\mathcal{A}$. Thus, the $\pi^n_{\mathcal{A},\vec{a}}$ are positive formulas. A theorem similar to Theorem 3.2 (p. 24) holds.

23♣ Formulate and prove a theorem that relates the appropriate version of the Ehrenfeucht game to R-*positive* sentences.

24♣ Modify the Ehrenfeucht game of length n on models $\mathcal{A}$ and $\mathcal{B}$ by requiring that Spoiler always picks her moves from A. Formulate and prove the corresponding modification of Theorem 2.4.

3.3 Infinite Game

3.18 Definition. In the *infinite* Ehrenfeucht game on $\mathcal{A}$ and $\mathcal{B}$, there is no bound on the number of moves; Spoiler and Duplicator alternate in making an ω-sequence of moves each, and win and loss are determined (almost) as before: Duplicator *wins* if at each finite stage of the play, the moves made so far constitute a local isomorphism between the models.
$\mathcal{A}$ and $\mathcal{B}$ are *partially isomorphic* if Duplicator has a winning strategy for the infinite game on $\mathcal{A}$ and $\mathcal{B}$.

3.19 Examples.

1. η and λ are partially isomorphic. Better still:
2. Every two dense linear orderings without endpoints are partially isomorphic.
3. No well-ordering is partially isomorphic with a non-well-ordering. (Let Spoiler play an infinite descending sequence in the non-well-ordering. Note that this argument also works for the 2-pebble game.)
4. Well-Orderings of different type are not partially isomorphic. (To begin with, Spoiler plays the element a of the larger one such that $a{\downarrow}$ has the type of the smaller one. Subsequently, Spoiler can always counter a move b of Duplicator with a move c such that $c{\downarrow}$ and $b{\downarrow}$ have the same type. Eventually, she must out-play Duplicator. For this argument, again 2 pebbles suffice.)

3.20 Determinacy. *In every infinite Ehrenfeucht game, exactly one of the players has a winning strategy.*

Proof. Note that if Spoiler wins a play, this has become apparent after finitely many moves already: the game is *open*, and the result is an instance of the Gale-Stewart Theorem. The argument proceeds as follows.

Suppose that Spoiler doesn't have a winning strategy, i.e.: that the initial position is no *win* for Spoiler. The result follows from the

Claim. *Avoiding positions that are wins for* Spoiler *makes* Duplicator *win.*

To begin with, the initial position satisfies this condition by assumption. Also, this happens to be a condition that Duplicator is able to preserve (i.e., this is a *strategy* for Duplicator): suppose that h is no win for Spoiler, and Spoiler plays, say, an element a in the first model. If, for every b in the other model, $h \cup \{(a, b)\}$ is a win for Spoiler, then a was a winning move for Spoiler and h would've been a win for Spoiler to begin with, contrary to assumption. Thus, Duplicator has a move b that brings him to a position that, again, is no win for Spoiler.

Finally: this strategy for Duplicator is winning. For, suppose it isn't. Then some play in which Duplicator uses this strategy is won by Spoiler. But that Spoiler wins will show after finitely many rounds. The corresponding position is trivially a win for Spoiler, contradicting the fact that the strategy avoids such positions. ⊣

The following important theorem has an extremely simple proof.

3.21 Theorem. *Countable partially isomorphic models are isomorphic.*

Proof. If Spoiler enumerates all elements of the two models and Duplicator uses his winning strategy, the relation that is built up during the play constitutes an isomorphism as required. ⊣

Cantor's characterization of the ordering η of the rationals is an immediate corollary. The proof of Theorem 3.21 is an abstract version of the usual back-and-forth proof for the Cantor result.

3.22 Corollary. *The linear ordering η is (up to isomorphism) the only countable dense linear ordering without endpoints.*

Of course, the homomorphism game has an infinite version as well, with its corresponding notion of *partial homomorphic*. Theorem 3.21 now modifies to:

3.23 Theorem. *If the countable models $\mathcal{A}$ and $\mathcal{B}$ are partially homomorphic, then there is a (surjective) homomorphism from $\mathcal{A}$ onto $\mathcal{B}$.*

Similarly:

3.24 Theorem. *If* Duplicator *has a winning strategy in the infinite* R*-positive game on the countable* $L \cup \{\mathrm{R}\}$*-models* $\mathcal{A}$ *and* $\mathcal{B}$*, then* $\mathcal{A}|L \cong \mathcal{B}|L$ *and the isomorphism is an* R*-homomorphism.*

Explaining the logical meaning of the infinite game needs the notion of an *infinitary* formula. This is obtained by modifying the definition of first-order formula, admitting conjunctions and disjunctions of arbitrarily many formulas. I.e., if L is a vocabulary, the class $L_{\infty\omega}$ of *infinitary* L*-formulas* is obtained by allowing (next to the usual rules) the following rule of formula-formation:

if $\Phi \subseteq L_{\infty\omega}$ is a *set*, then $\bigwedge \Phi, \bigvee \Phi \in L_{\infty\omega}$.

(In this notation, the ∞ signifies that arbitrary conjunctions and disjunctions are admitted; the ω indicates that quantification still is restricted to finitely many variables at the same time.)

The semantics of such infinitary formulas is obvious: the formula $\bigwedge \Phi$ (resp., $\bigvee \Phi$) is satisfied by the assignment α in the model $\mathcal{A}$ iff *every* (resp., *some*) $\varphi \in \Phi$ is. (This implies that $\bigwedge \emptyset$ is always satisfied whereas $\bigvee \emptyset$ never is, and that $\bigwedge\{\varphi\}$ and $\bigvee\{\varphi\}$ are logically equivalent with φ.) Equivalence with respect to infinitary sentences is denoted by $\equiv_{\infty\omega}$.

The following proposition explains that the infinite game is not just the limit of the finite games.

Recall the monotone operator $\Gamma = \Gamma_{\mathcal{A},\mathcal{B}}$ from 3.6 (p. 25):

$$\Gamma(X) = \{h \mid \forall a \in A\, \exists b \in B\, (h \cup \{(a,b)\} \in X) \ \wedge\ \forall b \in B\, \exists a \in A\, (h \cup \{(a,b)\} \in X)\},$$

of which the finite stages $\Gamma\!\downarrow n$ in its downward hierarchy were relevant to the finite game (Theorem 3.12 p. 27). Let W be the set of relations $\{(a_1,b_1),\dots,(a_n,b_n)\}$ such that Duplicator has a winning strategy for the infinite game on $(\mathcal{A},a_1,\dots,a_n)$ and $(\mathcal{B},b_1,\dots,b_n)$. Let EQ be the set of relations $\{(a_1,b_1),\dots,(a_n,b_n)\}$ such that $(\mathcal{A},a_1,\dots,a_n) \equiv_{\infty\omega} (\mathcal{B},b_1,\dots,b_n)$.

The second equality of the following result generalizes the fact that Duplicator has a winning strategy for the infinite game between two models iff they cannot be distinguished using infinitary sentences.

3.25 Proposition. $\Gamma\!\downarrow = W = EQ$.

Proof. By Lemma 3.8 (p. 25), it suffices to show that both W and EQ are co-inductive post-fixed points.

W is a post-fixed point: trivial.

W is co-inductive: Assume that X is a set of local isomorphisms such that $X \subseteq \Gamma(X)$. Suppose that $h \in X$. To see that $h \in W$, consider the strategy of Duplicator to satisfy, for every position $\{(a_1,b_1),\dots,(a_n,b_n)\}$ visited in the playing of the game, that $h \cup \{(a_1,b_1),\dots,(a_n,b_n)\} \in X$. If Duplicator succeeds in preserving this condition, he wins. That he can succeed follows from X being a post-fixed point.

EQ is co-inductive: Assume that $X \subseteq \Gamma(X)$. It follows that every $h := \{(a_1,b_1),\dots,(a_n,b_n)\} \in X$ satisfies $(\mathcal{A},a_1,\dots,a_n) \equiv_{\infty\omega} (\mathcal{B},b_1,\dots,b_n)$ using induction on sentences (keeping h variable).

EQ is a post-fixed point: Assume that $h := \{(a_1, b_1), \ldots, (a_n, b_n)\} \in EQ$. To see that $h \in \Gamma(EQ)$, suppose $a \in A$; we need to find $b \in B$ such that $h \cup \{(a, b)\} \in EQ$. If such a b doesn't exist, this means that for every $b \in B$ there is an infinitary formula $\varphi_b(x)$ such that $(\mathcal{A}, a_1, \ldots, a_n) \models \varphi_b[a]$ and $(\mathcal{B}, b_1, \ldots, b_n) \models \neg\varphi_b[b]$. Thus, we have that $(\mathcal{A}, a_1, \ldots, a_n) \models \exists x \bigwedge_{b \in B} \varphi_b$ and $(\mathcal{B}, b_1, \ldots, b_n) \models \neg\exists x \bigwedge_{b \in B} \varphi_b$, contradicting $h \in EQ$. ⊣

25♣ Let C be a (countably) infinite set of constant symbols. Show that the infinitary sentence $\forall x \bigvee_{c \in C} x = c$ doesn't have a first-order equivalent.

26♣ Suppose that $\mathcal{A} = (A, <)$ is a well-ordering. Recursively define, for $a \in A$, the infinitary formula φ_a as $\forall y(y < x \leftrightarrow \bigvee_{b<a} \varphi_b(y))$. (If you encounter problems with substituting into an infinitary formula, you might use $\forall y(y < x \leftrightarrow \exists x(y = x \wedge \bigvee_{b<a} \varphi_b))$. Thus, every φ_a uses two variables x and y; exactly one occurrence of x is free.) Let $\Phi_{\mathcal{A}}$ be the infinitary sentence $\forall x \bigvee_{a \in A} \varphi_a \wedge \bigwedge_{a \in A} \exists x \varphi_a$.

Show:

1. $(A, <) \models \varphi_a[b]$ iff $b = a$,
2. a linear ordering satisfies $\Phi_{\mathcal{A}}$ iff it is an isomorph of $\mathcal{A}$.

Note the straightforward generalization for models $\mathcal{A} = (A, \in)$ with A a transitive set.

Bisimulations. Suppose that the vocabulary L consists of some unary relation symbols plus one binary relation symbol R. Modify the formula formation rules for $L_{\infty\omega}$ by allowing only non-R-atoms in the one variable x and R-*bounded quantification*; that is: replace the quantification rules by:

If $\varphi = \varphi(x) \in L_{\infty\omega}$,
then $\forall y(\mathrm{R}(x, y) \to \varphi(y)), \exists y(\mathrm{R}(x, y) \wedge \varphi(y)) \in L_{\infty\omega}$.

3.26 Theorem. *For any two L-models $\mathcal{A}$ and $\mathcal{B}$ and elements $a \in A$ and $b \in B$, the following are equivalent:*

1. Duplicator *has a winning strategy for the infinite pebble game on $\mathcal{A}$ and $\mathcal{B}$ with just one pair of pebbles, starting at the initial position (a, b), where the moves are "R-restricted",*
2. *there is a bisimilation between $\mathcal{A}$ and $\mathcal{B}$ containing (a, b),*
3. *a and b satisfy the same (modified) $L_{\infty\omega}$-formulas in $\mathcal{A}$, resp., $\mathcal{B}$.*

Proof. A bisimulation is the same as a non-empty post-fixed point for the operator associated with the infinite pebble game. ⊣

Finitizing. The quantifier rank of an infinitary formula is defined by stipulating that, for the infinitary connectives:

$$\mathrm{qr}(\bigwedge \Phi) = \mathrm{qr}(\bigvee \Phi) = \sup\{\mathrm{qr}(\varphi) + 1 \mid \varphi \in \Phi\}.$$

The α-game is the modification of the infinite game in which Spoiler is required to choose, together with her moves, a descending sequence of ordinals $< \alpha$. A play of the game ends as soon as Duplicator has countered the move of Spoiler that goes with the ordinal 0. Thus, every play has finite length. For $\alpha = n < \omega$, the α-game is the same as the ordinary length-n Ehrenfeucht game. For $\alpha \geqslant \omega$, Spoiler can make a play of the α-game last as long as she wishes (in order to have better chances to win). E.g., if Spoiler chooses a finite ordinal $n \in \omega$ to begin with, the remaining game has $\leqslant n$ rounds. If $\alpha > \omega$ and Spoiler starts by chosing ω, she postpones the decision how long the play will be for her second move. Chosing $\omega + \omega + 1$ is a promise to tell, ultimately at the 3rd move, how long she will keep postponing the decision about the length, etc.

3.27 Theorem. *The following conditions are equivalent:*

1. Duplicator *has a winning strategy for the α-game,*
2. $\emptyset \in \Gamma{\downarrow}\alpha$,
3. *$\mathcal{A}$ and $\mathcal{B}$ have the same true rank $< \alpha$-sentences.*

The closure ordinal of $\Gamma_{\mathcal{A},\mathcal{A}}$ is called the *Scott rank* of $\mathcal{A}$.

27♣ Show that the Scott rank of the linear ordering ω equals ω.
Give an example of a model with Scott rank $> \omega$.

3.28 Characteristics. Recall Definition 3.1 (p. 23). Let $\mathcal{A}$ be a model. For every finite sequence $\vec{a} = (a_1, \ldots, a_n)$ from A and every ordinal α, the infinitary quantifier rank-α formula $\varepsilon^{\alpha}_{\mathcal{A},\vec{a}}(x_0, \ldots, x_{k-1})$ is defined, using recursion w.r.t. α, essentially as before, using infinite conjunctions to get across limit ordinals γ:

$$\varepsilon^{\gamma}_{\mathcal{A},\vec{a}}(x_0, \ldots, x_{k-1}) = \bigwedge_{\xi<\gamma} \varepsilon^{\xi}_{\mathcal{A},\vec{a}}(x_0, \ldots, x_{k-1}).$$

3.29 Theorem. Again we have:

1. for all α, $\mathcal{A} \models \varepsilon^{\alpha}_{\mathcal{A},\vec{a}}[\vec{a}]$,
2. $\mathcal{B} \models \varepsilon^{\alpha}_{\mathcal{A},\vec{a}}[\vec{b}]$

   iff $(\mathcal{A}, \vec{a})$ and $(\mathcal{B}, \vec{b})$ satisfy the same quantifier rank $\leq \alpha$ formulas,

   iff $\varepsilon^{\alpha}_{\mathcal{B},\vec{b}} = \varepsilon^{\alpha}_{\mathcal{A},\vec{a}}$.

If α is the Scott rank of $\mathcal{A}$, then $\varepsilon^{\alpha}_{\mathcal{A},\emptyset} \wedge \bigwedge_{\vec{a}} \forall \vec{x}(\varepsilon^{\alpha}_{\mathcal{A},\vec{a}} \to \varepsilon^{\alpha+1}_{\mathcal{A},\vec{a}})$ is the *Scott sentence* of $\mathcal{A}$.

The language $L_{\omega_1\omega}$ is the restriction of $L_{\infty\omega}$ that allows conjunctions and disjunctions over countable sets of formulas only.

Note that the Scott sentence of a countable model belongs to this language.

28♣ Show that the Scott sentence of a model axiomatizes its infinitary theory.

Summing Up. The different characterisations pertaining to the finite, infinite, and finitized game are collected.

For the finite game, the following are equivalent:

1. $\mathcal{A} \equiv^n \mathcal{B}$,
2. $\mathsf{D}(\mathcal{A}, \mathcal{B}, n)$,
3. $\mathcal{B} \models \varepsilon^n_{\mathcal{A}}$,
4. there is a length-$(n+1)$ Fraïssé-sequence,
5. $\emptyset \in \Gamma_{\mathcal{A},\mathcal{B}}\!\downarrow n$.

For the infinite game, this list becomes:

1. $\mathcal{A} \equiv_{\infty\omega} \mathcal{B}$,
2. $\mathsf{D}(\mathcal{A}, \mathcal{B}, \infty)$,
3. for all α, $\mathcal{B} \models \varepsilon^\alpha_{\mathcal{A}}$,
   equivalently: $\mathcal{B}$ satisfies the Scott sentence of $\mathcal{A}$,
4. there is a Fraïssé-sequence of type $\omega^\star$ (cf. Exercise 29);
   equivalently: there is a *partial isomorphism* between $\mathcal{A}$ and $\mathcal{B}$, i.e., a non-empty set I of local isomorphisms satisfying the back-and-forth condition (in other words: I is a non-empty post-fixedpoint of $\Gamma_{\mathcal{A},\mathcal{B}}$),
5. $\emptyset \in \Gamma_{\mathcal{A},\mathcal{B}}\!\downarrow$.

Finally, for the finitized version using ordinals, we have the following pairwise equivalent statements:

1. $\mathcal{A} \equiv^\alpha_{\infty\omega} \mathcal{B}$,
2. $\mathsf{D}(\mathcal{A}, \mathcal{B}, \alpha)$,
3. $\mathcal{B} \models \varepsilon^\alpha_{\mathcal{A}}$,
4. there is a length-$(\alpha+1)$ (Fraïssé-) Karp sequence,
5. $\emptyset \in \Gamma_{\mathcal{A},\mathcal{B}}\!\downarrow \alpha$.

29♣ A *Fraïssé-sequence of type* $\omega^\star$ for $\mathcal{A}$ and $\mathcal{B}$ is a sequence $\ldots, I_2, I_1, I_0$ of non-empty sets of local isomorphisms such that for all n and $h \in I_n$:

$\forall a \in A \exists b \in B(h \cup \{(a,b)\} \in I_{n+1})$ and $\forall b \in B \exists a \in A(h \cup \{(a,b)\} \in I_{n+1})$.

A *partial isomorphism* between $\mathcal{A}$ and $\mathcal{B}$ is a non-empty set I of local isomorphisms such that for all $h \in I$:

$\forall a \in A \exists b \in B(h \cup \{(a,b)\} \in I)$ and $\forall b \in B \exists a \in A(h \cup \{(a,b)\} \in I)$.

Show:

1. If I is a partial isomorphism, then $\ldots, I, I, I$ is a Fraïssé-sequence of type $\omega^\star$,
2. If $\ldots, I_2, I_1, I_0$ is a Fraïssé-sequence of type $\omega^\star$, then $\bigcup_n I_n$ is a partial isomorphism.

3.4 Fixed Points and Games

Let $\Gamma : \wp(I) \to \wp(I)$ be a monotone operator over a set I, and $h \in I$. Consider the following 2-person game. Players are *Challenger* ($\mathcal{C}$) and *Defender* ($\mathcal{D}$).

$\mathcal{D}$ starts by picking some $H_0 \subseteq I$ such that $h \in \Gamma(H_0)$.

$\mathcal{C}$ chooses $h_0 \in H_0$.

$\mathcal{D}$ chooses $H_1 \subseteq I$ such that $h_0 \in \Gamma(H_1)$.

$\mathcal{C}$ chooses $h_1 \in H_1$.

etc.

If one of the players cannot move, the other one *wins*. I.e., if $\mathcal{D}$ is able to select $\emptyset$, he wins ($\mathcal{C}$ is unable to pick an element in $\emptyset$); if $\mathcal{C}$ is able to select some $h_i \notin \Gamma(I)$, she wins (if $X \subseteq I$, then $\Gamma(X) \subseteq \Gamma(I)$; hence, $h_i \notin \Gamma(X)$). A never-ending play of the game is won by noone.

3.30 Theorem.

1. *$\mathcal{D}$ has a winning strategy iff $h \in \Gamma\uparrow$,*
2. *$\mathcal{C}$ has a winning strategy iff $h \notin \Gamma\downarrow$.*

Proof. 1. Define

$$W := \{h \in I \mid \mathcal{D} \text{ has a winning strategy for the game that starts at } h\}.$$

To see that $W = \Gamma\uparrow$, we check the two crucial properties.

(a) (Pre-fixedpoint property.) $\Gamma(W) \subseteq W$: Assume $h \in \Gamma(W)$. In position h, $\mathcal{D}$ plays W, and clearly wins ($\mathcal{C}$ must pick an element in W which represents positions won by $\mathcal{D}$).

(b) (Induction.) Assume that $\Gamma(Y) \subseteq Y$. Want: $W \subseteq Y$. Thus, suppose $h \notin Y$. Claim: $\mathcal{C}$ has a strategy by which he cannot lose, nl.: always playing elements $\notin Y$. (And, hence, $h \notin W$.) For: if $h \in \Gamma(X)$, then $X - Y \neq \emptyset$. (Else $X \subseteq Y$, $\Gamma(X) \subseteq \Gamma(Y)$, $h \in \Gamma(Y) \subseteq Y$.)

[[Alternatively: Assume that $h \in W$. Fix a winning strategy for $\mathcal{D}$ in position h. When $\mathcal{D}$ uses this strategy, all plays of the game are finite and won by $\mathcal{D}$. Thus, the tree of these plays is well-founded, and we can induct on it. Suppose that H_0 is the answer to h as given by this strategy. I.e., $h \in \Gamma(H_0)$ and $H_0 \subseteq W$. By induction hypothesis, $H_0 \subseteq Y$. Thus, $h \in \Gamma(H_0) \subseteq \Gamma(Y) \subseteq Y$, and $h \in Y$.]]

2. Define

$$L := \{h \in I \mid \mathcal{C} \text{ does not have a winning strategy for the game that starts at } h\}.$$

To show that $\Gamma\downarrow = L$, again the two crucial properties are verified.

(a) (Post-fixedpoint property.) $L \subseteq \Gamma(L)$: Assume that $h \in L$, i.e., $\mathcal{C}$ has no winning strategy in h. Thus, $\mathcal{D}$ has a move H_0 such that $h \in \Gamma(H_0)$ and $H_0 \subseteq L$. Then $\Gamma(H_0) \subseteq \Gamma(L)$ and $h \in \Gamma(L)$.

(b) (Co-induction.) Assume that $Y \subseteq \Gamma(Y)$. Want: $Y \subseteq L$. Let $h \in Y$. Obviously, $\mathcal{D}$ can repeat playing Y ad infinitum and, in doing so, demonstrates that $\mathcal{C}$ cannot have a winning strategy. Thus, $h \in L$. $\dashv$

If I is the set of local isomorphisms between two models $\mathcal{A}$ and $\mathcal{B}$ and $\Gamma = \Gamma_{\mathcal{A},\mathcal{B}}$ is the Ehrenfeucht operator, the game described above is reminiscent of the ordinary Ehrenfeucht game, with $\mathcal{C}$ playing the role of Spoiler and $\mathcal{D}$ that of Duplicator. The difference is, that here, in position $h \in I$, a move H by $\mathcal{D}$ such that $h \in \Gamma(H)$ comprises in a sense all answers of Duplicator on moves of Spoiler, and a next choice by $\mathcal{C}$ of $h' = h \cup \{(a,b)\} \in H$ compares to choosing by Spoiler one of the possibilities that Duplicator ($\mathcal{D}$) is offering with H.

Note that, in this particular case, no move of $\mathcal{D}$ can be empty. Thus, a play can be finite only if some position $g \in I$ is reached in which $\mathcal{D}$ cannot move, i.e., for which no H exists such that $g \in \Gamma(H)$.

Taking this parallel seriously, it is better to redefine *winning* so that $\mathcal{D}$ wins the *infinite* plays of the game, and $\mathcal{C}$ the *finite* ones. For that case, the above Theorem modifies to the simpler:

3.31 Theorem. *$\mathcal{D}$ has a winning strategy in h iff $h \in \Gamma{\downarrow}$.*

This shows that, indeed, the parallel is correct, since also Duplicator has a win in h iff $h \in \Gamma{\downarrow}$.

Proof. $\Leftarrow$ If $h \in \Gamma{\downarrow}$ $(= \Gamma(\Gamma{\downarrow}))$, $\mathcal{D}$ persists in repeating $\Gamma{\downarrow}$ and wins.

$\Rightarrow$ The set $W = \{h \mid h \text{ is a win for } \mathcal{D}\}$ is a post-fixed point of Γ (and, hence, included in $\Gamma{\downarrow}$): Suppose that h is a win for $\mathcal{D}$. Then $\mathcal{D}$ has a winning move, that is: there is some H such that $h \in \Gamma(H)$ and $H \subseteq W$. But then $\Gamma(H) \subseteq \Gamma(W)$, and $h \in \Gamma(W)$. $\dashv$

Chapter 4

Applications involving Compactness

This part exploits characteristics together with compactness to prove several classics of first-order logic. It is inspired by Barwise and van Benthem: *Interpolation, preservation and pebble games*, JSL 64 (1999) 881–903 (modified for the first-order setting); it shows that characteristics provide a tool with which all these results can be obtained in a uniform way.

4.1 Interpolation & Co.

The following Lemma constitutes the basic trick involving Compactness that is used below.

The proof uses (expansions of) *model pairs*, a construct that can be implemented in several ways.

Model Pairs. Suppose that $\mathcal{A}_i$ is an L_i-model ($i = 1, 2$) and L is the disjoint union of L_1 and L_2 together with two new unary relation symbols $\mathrm{S}_1, \mathrm{S}_2$. The *model pair* $\mathcal{A} = (\mathcal{A}_1, \mathcal{A}_2)$ is the L-model with universe $A_1 \cup A_2$, with $\mathrm{S}_i^{\mathcal{A}} = A_i$ ($i = 1, 2$), and where the symbols of the L_i retain their old meanings.

4.1 Lemma. *Suppose that $L = L_1 \cap L_2$, and that $\mathcal{A}_1$ and $\mathcal{A}_2$ are L_1, resp., L_2-models such that $\mathcal{A}_1|L \equiv \mathcal{A}_2|L$. Then countable models $\mathcal{B}_i \equiv \mathcal{A}_i$ exist ($i = 1, 2$) such that $\mathcal{B}_1|L \cong \mathcal{B}_2|L$.*

Proof. Assume that $\mathcal{A}_1|L \equiv \mathcal{A}_2|L$. By Remark 3.5 (p. 25), for every $n \in \mathbb{N}$ there is a sequence $R_0 = \top, \dots, R_n$ of relations $R_i \subseteq A_1^i \times A_2^i$, satisfying

1. for $0 \leqslant i \leqslant n$, if $R_i(a_1, \dots, a_i, b_1, \dots, b_i)$, then $\{(a_1, b_1), \dots, (a_i, b_i)\}$ is a local isomorphism between $\mathcal{A}_1|L$ and $\mathcal{A}_2|L$, and
2. for $0 \leqslant i < n$, if $R_i(a_1, \dots, a_i, b_1, \dots, b_i)$, then both
   - $\forall a \in A_1\ \exists b \in A_2\ R_{i+1}(a_1, \dots, a_i, a, b_1, \dots, b_i, b)$, and
   - $\forall b \in A_2\ \exists a \in A_1\ R_{i+1}(a_1, \dots, a_i, a, b_1, \dots, b_i, b)$.

These conditions on complex models of the form $(\mathcal{A}_1, \mathcal{A}_2, R_0, R_1, R_2, \dots)$ (expansions of model pairs) can be formulated in first-order terms, using new relation symbols $\mathrm{R}_0, \mathrm{R}_1, \mathrm{R}_2, \dots$. Thus, by Compactness and Downward Löwenheim-Skolem, there is a countable model $(\mathcal{B}_1, \mathcal{B}_2, S_0, S_1, S_2, \dots)$ satisfying $\mathcal{B}_i \equiv \mathcal{A}_i$ $(i = 1, 2)$ and such that the sequence $S_0, S_1, S_2, \dots$ obeys the above conditions w.r.t. $\mathcal{B}_1, \mathcal{B}_2$ for *all* i.

It follows that the associated set

$$\bigcup_i \{\{(a_1, b_1), \dots, (a_i, b_i)\} \mid S_i(a_1, \dots, a_i, b_1, \dots, b_i)\}$$

of local isomorphisms between $\mathcal{B}_1|L$ and $\mathcal{B}_2|L$ is a non-empty post-fixed point for the relevant game operator $\Gamma = \Gamma_{\mathcal{B}_1|L, \mathcal{B}_2|L}$. Thus, $\Gamma\downarrow$ is non-empty, the models $\mathcal{B}_1|L$ and $\mathcal{B}_2|L$ are partially isomorphic, and hence (by Theorem 3.21 p. 30), isomorphic. ⊣

4.2 Consistency Theorem (Robinson). *Suppose that T_i is a set of L_i-sentences $(i = 1, 2)$ such that $T_1 \cup T_2$ has no model. Then there is an L-sentence φ (where $L = L_1 \cap L_2$) such that $T_1 \models \varphi$ and $T_2 \models \neg\varphi$.*

Proof. Suppose that no such φ exists. The following constructs a model for $T_1 \cup T_2$.

Claim. *For all n, there exist $\mathcal{A} \models T_1$ and $\mathcal{B} \models T_2$ such that $\mathcal{B} \models \varepsilon^n_{\mathcal{A}|L}$.*

Proof. Suppose this fails for the integer n. Consider the finite set of L-sentences $\Sigma = \{\varepsilon^n_{\mathcal{A}|L} \mid \mathcal{A} \models T_1\}$. Put $\varphi = \bigvee \Sigma$. It suffices to show that both $T_1 \models \varphi$ and $T_2 \models \neg\varphi$.

As to the first statement, assume that $\mathcal{A} \models T_1$. Then $\mathcal{A} \models \varepsilon^n_{\mathcal{A}|L} \in \Sigma$, and hence $T_1 \models \bigvee \Sigma$.

As to the second one, assume that $\mathcal{B} \models T_2$ and $\mathcal{B} \models \bigvee \Sigma$. Then for some $\mathcal{A} \models T_1$ we have that $\mathcal{B} \models \varepsilon^n_{\mathcal{A}|L}$, contradicting the assumption on n. ⊣

Applying Compactness to this Claim, we obtain $\mathcal{A} \models T_1$ and $\mathcal{B} \models T_2$ such that $\mathcal{A}|L \equiv \mathcal{B}|L$. Applying Lemma 4.1, we obtain (countable) $\mathcal{A} \models T_1$ and $\mathcal{B} \models T_2$ such that $\mathcal{A}|L \cong \mathcal{B}|L$. Identifying $\mathcal{A}|L$ and $\mathcal{B}|L$ results in a model for $T_1 \cup T_2$. ⊣

4.3 Interpolation Theorem (Craig). *Suppose that $L = L_1 \cap L_2$, and the sentences $\varphi_i \in L_i$ $(i = 1, 2)$ are such that $\varphi_1 \models \varphi_2$. Then a sentence $\varphi \in L$ (an* interpolant*) exists such that both $\varphi_1 \models \varphi$ and $\varphi \models \varphi_2$.*

Proof. There is a standard easy argument using the Consistency Theorem (taking $T_1 = \{\varphi_1\}$ and $T_2 = \{\neg\varphi_2\}$). However, since we also want to deal with Lyndon's refinement below, here follows the straightforward proof in the style of the above one.

Suppose there is no interpolant.

Claim. *For every $n \in \mathbb{N}$ there exist $\mathcal{A} \models \varphi_1$ and $\mathcal{B} \models \neg\varphi_2$ such that $\mathcal{B} \models \varepsilon^n_{\mathcal{A}|L}$.*

Proof. If this happens to be false for n, consider the set $\Sigma = \{\varepsilon^n_{\mathcal{A}|L} \mid \mathcal{A} \models \varphi_1\}$. Note that Σ is a finite set of L-sentences. We claim that $\bigvee \Sigma$ is an interpolant. Indeed: if $\mathcal{A} \models \varphi_1$, then $\mathcal{A} \models \varepsilon^n_{\mathcal{A}|L} \in \Sigma$, and, hence, $\mathcal{A} \models \bigvee \Sigma$. And if $\mathcal{B} \models \bigvee \Sigma$, say, $\mathcal{B} \models \varepsilon^n_{\mathcal{A}|L}$, where $\mathcal{A} \models \varphi_1$, then, by assumption on n, $\mathcal{B} \models \varphi_2$. ⊣

As in the proof of the Consistency Theorem, Lemma 4.1 can now be applied to yield a counter-model to $\varphi_1 \models \varphi_2$. ⊣

30♣ Cf. Theorem 4.3. Suppose that $\varphi_i \in L_i$ $(i = 1, 2)$ and $L = L_1 \cap L_2$. Show that the following are equivalent:

1. $\forall \mathcal{A}, \mathcal{B}(\mathcal{A} \models \varphi_1 \ \wedge \ \mathcal{A}|L = \mathcal{B}|L \ \Rightarrow \ \mathcal{B} \models \varphi_2)$,
2. $\forall \mathcal{A}, \mathcal{B}(\mathcal{A} \models \varphi_1 \ \wedge \ \mathcal{A}|L \equiv_{\infty\omega} \mathcal{B}|L \ \Rightarrow \ \mathcal{B} \models \varphi_2)$,
3. $\forall \mathcal{A}, \mathcal{B}(\mathcal{A} \models \varphi_1 \ \wedge \ \mathcal{A}|L \equiv \mathcal{B}|L \ \Rightarrow \ \mathcal{B} \models \varphi_2)$,
4. $\exists n \forall \mathcal{A}, \mathcal{B}(\mathcal{A} \models \varphi_1 \ \wedge \ \mathcal{A}|L \equiv^n \mathcal{B}|L \ \Rightarrow \ \mathcal{B} \models \varphi_2)$.

4.4 Refinement of the Interpolation Theorem (Lyndon). *Same as 4.3, but the interpolant for $\varphi_1 \models \varphi_2$ has to satisfy additional polarity requirements: relation symbols (different from =) occurring positively (resp., negatively) in the interpolant should occur positively (resp., negatively) in both φ_1 and φ_2.*

Proof. Modify the argument for 4.3 as follows.

First, let P_i be the set of relation symbols $\mathrm{R} \in L = L_1 \cap L_2$ that occur positively in φ_i $(i = 1, 2)$ and let N_i be the set of $\mathrm{R} \in L$ that occur negatively in φ_i.

Define the $\varepsilon^n_{\mathcal{A},\vec{a}}$ for $n > 0$ as before, but, this time, let $\varepsilon^0_{\mathcal{A},\vec{a}}$ be the conjunction of (i) all atoms satisfied by $\vec{a}$ in $\mathcal{A}$ that carry a relation symbol in $P_1 \cap P_2$, (ii) all negated atoms satisfied by $\vec{a}$ in $\mathcal{A}$ that carry a relation symbol in $N_1 \cap N_2$, (iii) all =-atoms and negated =-atoms satisfied by $\vec{a}$ in $\mathcal{A}$. These are the obvious modifications to make if one wants to preserve the (proof for the) Claim in the proof of 4.3, since now, the interpolant $\bigvee \Sigma$ has to satify the additional polarity requirements.

Note that these modified characteristics are adequate w.r.t. the (asymmetric) Ehrenfeucht game on $\mathcal{A}$ and $\mathcal{B}$ in which the winning condition is changed to: Duplicator has won in the terminal position $\{(a_1, b_1), \ldots, (a_n, b_n)\}$ iff (i) $\mathrm{R}^{\mathcal{A}}(\vec{a}) \Rightarrow \mathrm{R}^{\mathcal{B}}(\vec{b})$ for $\mathrm{R} \in P_1 \cap P_2$, (ii) $\mathrm{R}^{\mathcal{B}}(\vec{b}) \Rightarrow \mathrm{R}^{\mathcal{A}}(\vec{a})$ for $\mathrm{R} \in N_1 \cap N_2$, (iii) the correspondence $a_i \leftrightarrow b_i$ is one-one. That is, Theorem 3.2 (p. 24) (: Duplicator has a winning strategy for the length-n game iff $\mathcal{B} \models \varepsilon^n_{\mathcal{A}}$) is literally true under these modifications.

The condition $\mathcal{B} \models \varepsilon^n_{\mathcal{A}|L}$ means that a sequence $R_0 = \top, \ldots, R_n$ of relations $R_i \subseteq A_i \times B_i$ exists satisfying the usual back-and-forth-conditions; however, the modifications (i) and (ii) above entail that they do *not* need to code sets of local isomorphisms between $\mathcal{A}|L$ and $\mathcal{B}|L$. What *does* hold is indicated in (i)–(iii) above.

After applying Compactness and Löwenheim-Skolem, we obtain a relation $h \subseteq A \times B$, generated by the (proof of) Theorem 3.21. From the way in which

h is constructed, we'll have that h is a bijection between $\mathcal{A}$ and $\mathcal{B}$. But again, for $\mathrm{R} \in L$, we'll only have that $\mathrm{R}^{\mathcal{A}}(a) \Rightarrow \mathrm{R}^{\mathcal{B}}(h(a))$ when $\mathrm{R} \in P_1 \cap P_2$, and $\neg\mathrm{R}^{\mathcal{A}}(a) \Rightarrow \neg\mathrm{R}^{\mathcal{B}}(h(a))$, i.e.: $\mathrm{R}^{\mathcal{B}}(h(a)) \Rightarrow \mathrm{R}^{\mathcal{A}}(a)$, when $\mathrm{R} \in N_1 \cap N_2$. For symbols in $P_1 \cap P_2 \cap N_1 \cap N_2$, h preserves in both directions (this is item 7 in the list below). However, for the remaining symbols, there is no preservation by h in any direction (items 5 and 6).

For the rest of the argument to make sense, however, we need not only that $\mathcal{A} \models \varphi_1$ and $\mathcal{B} \models \neg\varphi_2$, but also that $\mathcal{A}|L \cong \mathcal{B}|L$, for we have to form one model for $\varphi_1 \wedge \neg\varphi_2$ out of $\mathcal{A}$ and $\mathcal{B}$ by identifying $\mathcal{A}|L$ and $\mathcal{B}|L$.

The solution is to modify, in all cases but one, either the interpretation $\mathrm{R}^{\mathcal{A}}$ or $\mathrm{R}^{\mathcal{B}}$ of a symbol $\mathrm{R} \in L$, forcing h to be an isomorphism for the modifications, but preserving φ_1 in the modified $\mathcal{A}$ and $\neg\varphi_2$ in the modified $\mathcal{B}$.

Note: w.r.t. φ_1, a relation symbol $\mathrm{R} \in L$ can occur either in $P_1 - N_1$, in $N_1 - P_1$, or in $P_1 \cap N_1$; and the same goes for φ_2. Hence, all in all, there are $3 \times 3 = 9$ cases to be looked into. In the following list, both 1 and 2 consider two cases each, 3–7 consider one case each.

1. Replace $\mathrm{R}^{\mathcal{A}}$ by $R(a) :\equiv \mathrm{R}^{\mathcal{B}}(h(a))$ if R occurs in $P_1 - N_1$ and P_2.

   Note: since, in this case, $\mathrm{R}^{\mathcal{A}}(a) \Rightarrow \mathrm{R}^{\mathcal{B}}(h(a))$ holds, we have that $\mathrm{R}^{\mathcal{A}} \subseteq R$; and since $\mathrm{R} \in P_1 - N_1$, φ_1 will still hold in the modified $\mathcal{A}$.

2. Replace $\mathrm{R}^{\mathcal{A}}$ by $R(a) :\equiv \mathrm{R}^{\mathcal{B}}(h(a))$ if R occurs in $N_1 - P_1$ and N_2.

   Note: this time, $R \subseteq \mathrm{R}^{\mathcal{A}}$, and since $\mathrm{R} \in N_1 - P_1$, this preserves φ_1.

3. Enlarge $\mathrm{R}^{\mathcal{B}}$ to $R(b) :\equiv \mathrm{R}^{\mathcal{A}}(h^{-1}(b))$ if R occurs in $P_1 \cap N_1$ and $N_2 - P_2$.

   Note that $\neg\varphi_2$ still will hold in the modified $\mathcal{B}$.

4. Replace $\mathrm{R}^{\mathcal{B}}$ by the smaller $R(b) :\equiv \mathrm{R}^{\mathcal{A}}(h^{-1}(b))$ if R occurs in $P_1 \cap N_1$ and $P_2 - N_2$.

5. Replace both $\mathrm{R}^{\mathcal{A}}$ and $\mathrm{R}^{\mathcal{B}}$ by $\perp$ ("false") if R occurs in $N_1 - P_1$ and $P_2 - N_2$.

6. Replace both $\mathrm{R}^{\mathcal{A}}$ and $\mathrm{R}^{\mathcal{B}}$ by $\top$ ("true") if R occurs in $P_1 - N_1$ and $N_2 - P_2$.

7. In the remaining case, where R occurs in both $P_1 \cap N_1$ and $P_2 \cap N_2$, no relation has to be changed as preservation by h is already guaranteed.

$\dashv$

4.5 Definability Theorem (Beth). *Suppose that $L^+ = L \cup \{\mathrm{R}\}$ and that T is an L^+-theory such that for every two models $\mathcal{A}$ and $\mathcal{B}$ of T, if $\mathcal{A}|L = \mathcal{B}|L$, then $\mathrm{R}^{\mathcal{A}} = \mathrm{R}^{\mathcal{B}}$. Then an L-formula $\varphi = \varphi(x)$ (a* definition *of* R *w.r.t. T) exists such that*

$$T \models \forall x(\mathrm{R}(x) \leftrightarrow \varphi).$$

Proof. Again, there is a standard argument using Interpolation. However, here follows a direct one using characteristics.

Suppose no definition exists. We shall construct $\mathcal{A}, \mathcal{B} \models T$ such that $\mathcal{A}|L = \mathcal{B}|L$ but $\mathrm{R}^{\mathcal{A}} \neq \mathrm{R}^{\mathcal{B}}$.

Claim. *For all n there exist $\mathcal{A} \models T$ and $a \in A$ with $\mathrm{R}^{\mathcal{A}}(a)$, and $\mathcal{B} \models T$ and $b \in B$ with $\neg\mathrm{R}^{\mathcal{B}}(b)$, such that $\mathcal{B} \models \varepsilon^n_{\mathcal{A}|L,a}[b]$.*

Proof. Suppose this fails for n. Consider $\Sigma = \{\varepsilon^n_{\mathcal{A}|L,a} \mid \mathcal{A} \models T \wedge \mathrm{R}^{\mathcal{A}}(a)\}$. Then $\bigvee \Sigma$ is a definition for R: First, if $\mathcal{A} \models T$ and $\mathrm{R}^{\mathcal{A}}(a)$, then (since $\mathcal{A} \models \varepsilon^n_{\mathcal{A}|L,a}[a]$), we have that $\mathcal{A} \models \bigvee \Sigma[a]$. Second, if $\mathcal{B} \models T$ and $\mathcal{B} \models \bigvee \Sigma[b]$; say, $\mathcal{B} \models \varepsilon^n_{\mathcal{A}|L,a}[b]$, where $\mathcal{A} \models T$ and $\mathrm{R}^{\mathcal{A}}(a)$, then $\mathrm{R}^{\mathcal{B}}(b)$ holds by assumption on n. ⊣

The rest of the proof is as usual, using Lemma 4.1. Note that we can take care that, for the resulting (countable) models $\mathcal{A}$ and $\mathcal{B}$ and the isomorphism h between $\mathcal{A}|L$ and $\mathcal{B}|L$, there is $a \in A$ such that $\mathrm{R}^{\mathcal{A}}(a)$ but $\neg\mathrm{R}^{\mathcal{B}}(h(a))$. ⊣

4.2 Preservation under Homomorphism

A sentence is *preserved under homomorphisms* if it is true of every homomorphic image of one of its models.

4.6 Lyndon's Theorem. *A sentence is preserved under homomorphisms iff it has a positive logical equivalent.*

Proof. For one direction, positive sentences are easily seen to be preserved. For the other, more difficult one, we use the familiar argument, this time using the characteristics $\pi^n_{\mathcal{A}}$ relative to the homomorphism game (see p. 29).

Suppose that Φ has no positive equivalent.

Claim. *For every $n \in \mathbb{N}$ there exist $\mathcal{A} \models \Phi$ and $\mathcal{B} \models \neg\Phi$ such that $\mathcal{B} \models \pi^n_{\mathcal{A}}$.*

Proof. If this happens to be false for n, consider the set $\Pi = \{\pi^n_{\mathcal{A}} \mid \mathcal{A} \models \Phi\}$. Note that Π is a finite set of positive sentences. We claim that $\bigvee \Pi$ is a first-order equivalent of Φ. Indeed: if $\mathcal{A} \models \Phi$, then $\mathcal{A} \models \pi^n_{\mathcal{A}} \in \Pi$, and, hence, $\mathcal{A} \models \bigvee \Pi$. Conversely, if $\mathcal{B} \models \bigvee \Pi$, say, $\mathcal{B} \models \pi^n_{\mathcal{A}}$, where $\mathcal{A} \models \Phi$, then, by assumption on n, $\mathcal{B} \models \Phi$. ⊣

The rest of the proof follows the by now familiar pattern. The condition that $\mathcal{B} \models \pi^n_{\mathcal{A}}$ can be rewritten (In a way similar to Remark 3.5 p. 25) as the existence of a finite sequence of relations $R_0 = \top, \ldots, R_n$ coding sets of local homomorphisms that satisfy the usual back-and-forth properties.

By Downward Löwenheim-Skolem and Compactness, we obtain a countable complex $(\mathcal{A}, \mathcal{B}, R_0, R_1, R_2, \ldots)$ with $\mathcal{A} \models \Phi$, $\mathcal{B} \models \neg\Phi$, and such that

$$\bigcup_i \{\{(a_1, b_1), \ldots, (a_i, b_i)\} \mid R_i(a_1, \ldots, a_i, b_1, \ldots, b_i)\}$$

is a post-fixed point for the relevant game operator. Thus, $\mathcal{A}$ and $\mathcal{B}$ are partially homomorphic; and it follows that $\mathcal{B}$ is a homomorphic image of $\mathcal{A}$. Thus, Φ isn't homomorphism-preserved. ⊣

31♣ (Łoś-Tarski) Show that a sentence is preserved under model-extensions iff it has an existential equivalent.
Hint. Modify the characteristics appropriately: $\varepsilon^0_{\mathcal{A},\vec{a}}$ is as before, but, this time, $\varepsilon^{n+1}_{\mathcal{A},(a_1,\ldots,a_k)}$ is $\bigwedge_{a\in A}\exists x_{k+1}\varepsilon^n_{\mathcal{A},(a_1,\ldots,a_k,a)}$. Note: this modification yields *existential* formulas.

32♣ An $L\cup\{\mathrm{R}\}$-sentence Φ is *preserved under* R-*extensions* if for every two models $\mathcal{A}$ and $\mathcal{B}$, if $\mathcal{A}\models\Phi$, $\mathcal{A}|L=\mathcal{B}|L$ and $\mathrm{R}^{\mathcal{A}}\subseteq\mathrm{R}^{\mathcal{B}}$, then $\mathcal{B}\models\Phi$.

Show: a sentence is preserved under R-extensions iff it has an R-positive equivalent.

4.3 Modal Logic

4.7 Theorem. (van Benthem) *If a first-order formula in one free variable is preserved under bisimulation, then it has a modal equivalent (that is: an equivalent that is the* standard translation *of some modal formula).*

Proof. The "modal" vocabulary has a binary "accessibility" relation symbol R plus unary relation symbols U_j $(j\in J)$. (Kripke) models are of the form $\mathcal{A}=(A,\mathrm{R}^{\mathcal{A}},\mathrm{U}^{\mathcal{A}}_j)_{j\in J}$. If $\varphi=\varphi(x)$ is a formula with one free variable, by $\mathcal{A},a\models\varphi$ we mean that $\mathcal{A}\models\varphi[a]$. For a model $\mathcal{A}$ and an element $a\in A$, define the formulas $\sigma^n_a=\sigma^n_{\mathcal{A},a}$ in one free variable as follows:

1. $\sigma^0_a(x)$ is the conjunction of all U_j-literals $\mathrm{U}_j(x)$ and $\neg\mathrm{U}_j(x)$ that are satisfied by a in $\mathcal{A}$.
2. $\sigma^{n+1}_a(x)=\bigwedge_{\mathrm{R}^{\mathcal{A}}(a,b)}\exists y(\mathrm{R}(x,y)\wedge\sigma^n_b(y))\wedge\forall y(\mathrm{R}(x,y)\rightarrow\bigvee_{\mathrm{R}^{\mathcal{A}}(a,b)}\sigma^n_b(y))$.

Note that these are all (standard translations of) modal formulas. Obviously, we have that $\mathcal{A},a\models\sigma^n_a$.

Suppose that the first-order formula $\Phi(x)$ is preserved under bisimulation, but has no modal equivalent.

Claim. For all n there are $\mathcal{A},a\models\Phi$ and $\mathcal{B},b\models\neg\Phi$ such that $\mathcal{B},b\models\sigma^n_{\mathcal{A},a}$.

Proof. If this is false for n, consider $\Sigma=\{\sigma^n_{\mathcal{A},a}(x)\mid\mathcal{A},a\models\Phi\}$; now $\bigvee\Sigma$ would be a modal equivalent for Φ:

If $\mathcal{A},a\models\Phi$, then $\sigma_{\mathcal{A},a}\in\Sigma$, and hence $\mathcal{A},a\models\bigvee\Sigma$ holds.

Conversely, if $\mathcal{B},b\models\bigvee\Sigma$, say, $\mathcal{B},b\models\sigma^n_{\mathcal{A},a}$ where $\mathcal{A},a\models\Phi$, then $\mathcal{B},b\models\Phi$ by assumption on n. $\dashv$

Claim. Suppose that $\mathcal{B},b\models\sigma^n_{\mathcal{A},a}$ holds. Then relations $R_0,\ldots,R_n\subseteq A\times B$ exist such that

1. if $R_i(u,v)$, then $\mathrm{U}^{\mathcal{A}}_j(u)\Leftrightarrow\mathrm{U}^{\mathcal{B}}_j(v)$ $(j\in J)$,
2. $R_0(a,b)$,
3. (forth) $i<n$, $R_i(u,v)$ and $\mathrm{R}^{\mathcal{A}}(u,u')$ imply $\exists v'\in B[\mathrm{R}^{\mathcal{B}}(v,v')\wedge R_{i+1}(u',v')]$, (back) similar.

Proof. Define $R_i(u,v) :\equiv \mathcal{B}, v \models \sigma^{n-i}_{\mathcal{A},u}$. ⊣

By Compactness we find $\mathcal{A}, a$; $\mathcal{B}, b$; and $R_0, R_1, R_2, \ldots \subseteq A \times B$ such that $\mathcal{A}, a \models \Phi$; $\mathcal{B}, b \models \neg\Phi$; and such that the conditions of the second claim are satisfied for all i. It follows that $\mathcal{A}, a$ and $\mathcal{B}, b$ are bisimilar, contradicting the assumption on Φ. ⊣

4.8 Interpolation. *If φ_1 and φ_2 are modal formulas such that $\varphi_1 \models \varphi_2$, then a modal formula φ (an interpolant) exists such that $\varphi_1 \models \varphi$, $\varphi \models \varphi_2$, and every relation symbol U_j in φ occurs in both φ_1 and φ_2.*

Proof. Suppose that $\varphi_1 \models \varphi_2$, but an interpolant doesn't exist.

Modify the definition of the $\sigma^n_{\mathcal{A},a}$ by allowing, in σ^0_a, *only* U_j-literals where U_j occurs in both φ_1 and φ_2.

Claim. For all n there are $\mathcal{A}, a \models \varphi_1$ and $\mathcal{B}, b \models \neg\varphi_2$ such that $\mathcal{B}, b \models \sigma^n_{\mathcal{A},a}$.

By Compactness, obtain $\mathcal{A}, a \models \varphi_1$, and $\mathcal{B}, b \models \neg\varphi_2$, such that $\mathcal{A}, a$ and $\mathcal{B}, b$ are bisimilar w.r.t. relations common to the two formulas.

Needed: amalgamation into one model... ⊣

4.9 Łoś-Tarski Theorem. If a modal formula is preserved under model extensions, it has an existential modal equivalent.

Proof. Modify the above definition of the σ^n by putting

$$\sigma^{n+1}_a(x) \;=\; \bigwedge_{\mathrm{R}^{\mathcal{A}}(a,b)} \exists y(\mathrm{R}(x,y) \;\wedge\; \sigma^n_b(y)).$$

Suppose that $\Phi(x)$ is modal and preserved under model extensions, but has no existential modal equivalent.

Claim. For all n there are $\mathcal{A}, a \models \Phi$ and $\mathcal{B}, b \models \neg\Phi$ such that $\mathcal{B}, b \models \sigma^n_{\mathcal{A},a}$.

By Compactness, obtain $\mathcal{A}, a \models \Phi$, $\mathcal{B}, b \models \neg\Phi$ and $R_0, R_1, R_2, \ldots \subseteq A \times B$, where now only the "forth" condition is satisfied.

It suffices to find a submodel $\mathcal{B}' \subseteq \mathcal{B}$ such that $\mathcal{B}', b$ bisimulates $\mathcal{A}, a$. Put $R = \bigcup_n R_n$. Define $B' = \bigcup_n B_n$, where $B_0 = \{b\}$, and B_{n+1} is the least set $\supseteq B_n$ such that

$$R(u,v),\ v \in B_n \text{ and } \mathrm{R}^{\mathcal{A}}(u,u') \text{ imply } \exists v' \in B_{n+1}\, R(u',v').$$

⊣

4.4 Lindström's Theorem

The same argument is used once more to prove *Lindström's Theorem*, which says that Löwenheim-Skolem and Compactness *characterize* first-order logic in the following sense:

4.10 Lindström's Theorem. *There is no logic that properly extends first-order logic and still satisfies the Downward Löwenheim-Skolem and Compactness Theorems.*

Before proving this, you should get explained what is meant here by "a logic" that "properly extends first-order logic".

4.11 Logic. A *logic* is a schema $\mathcal{Z}$ that associates to any vocabulary L a set $\mathcal{Z}(L)$ of *sentences* together with a truth-relation $\models$ (better: $\models_{\mathcal{Z}(L)}$) between L-models and sentences from $\mathcal{Z}(L)$ such that the following conditions hold:

1. Isomorphic models have the same true $\mathcal{Z}$-sentences.
2. Suppose that L extends the vocabulary (L_1, L_2) that goes with model pairs $(\mathcal{A}_1, \mathcal{A}_2)$ built from an L_1-model $\mathcal{A}_1$ and an L_2-model $\mathcal{A}_2$. Then for every $\Phi \in \mathcal{Z}(L_i)$ $(i = 1, 2)$ there must be a sentence $\Phi^i \in \mathcal{Z}(L)$ such that for all L_i-models $\mathcal{A}_i$ $(i = 1, 2)$, if $(\mathcal{A}_1, \mathcal{A}_2, \ldots)$ is an L-expansion of $(\mathcal{A}_1, \mathcal{A}_2)$, then:

   $$(\mathcal{A}_1, \mathcal{A}_2, \ldots) \models \Phi^i \;\Leftrightarrow\; \mathcal{A}_i \models \Phi \quad (i = 1, 2).$$

   I.e.: in the logic $\mathcal{Z}$ we can express, relative to $(\mathcal{A}_1, \mathcal{A}_2, \ldots)$, that Φ holds in one of the component-models.

Remark. First-order logic is, indeed, a logic in the sense of 4.11.

What it means for a logic to (properly) *extend* first-order logic, we leave mostly to the reader's imagination. One example is the logic obtained from first-order logic by adding a quantifier symbol F with the meaning that $\mathcal{A} \models \mathrm{F}x\varphi(x)$ holds iff $\{a \in A \mid \mathcal{A} \models \varphi[a]\}$ is finite. Note that, e.g., $\mathrm{F}x(x{=}x)$ and $\forall y \mathrm{F}x(x < y)$ have no first-order equivalent.

What we need for the proof is: closure under negation (with the usual meaning) and inclusion of all first-order sentences in the given vocabulary. We need the Downward Löwenheim-Skolem Theorem in the following form: every satisfiable set of sentences in a countable vocabulary has a countable model.

Proof of Lindström's Theorem. Let $\mathcal{Z}$ be a logic satisfying the conditions, L a finite vocabulary, and Φ an arbitrary sentence in $\mathcal{Z}(L)$. We are going to show that Φ is first-order, that is: has a first-order equivalent.

The proof is by contradiction. Thus, suppose that Φ is *not* first-order.

Claim. *For every* $n \in \mathbb{N}$ *there exist* $\mathcal{A} \models \Phi$ *and* $\mathcal{B} \models \neg\Phi$ *such that* $\mathcal{B} \models \varepsilon^n_{\mathcal{A}}$.

Proof. If this happens to be false for n, consider the set $\Sigma = \{\varepsilon^n_{\mathcal{A}} \mid \mathcal{A} \models \Phi\}$. Note that Σ is finite. We claim that $\bigvee \Sigma$ is a first-order equivalent of Φ. Indeed: if $\mathcal{A} \models \Phi$, then $\mathcal{A} \models \varepsilon^n_{\mathcal{A}} \in \Sigma$, and, hence, $\mathcal{A} \models \bigvee \Sigma$. Conversely, if $\mathcal{B} \models \bigvee \Sigma$, say, $\mathcal{B} \models \varepsilon^n_{\mathcal{A}}$, $\mathcal{A} \models \Phi$, then, by assumption on n, $\mathcal{B} \models \Phi$. $\dashv$

The proof is finished in the usual way, using Lemma 4.1, by constructing (countable) $\mathcal{A} \models \Phi$ and $\mathcal{B} \models \neg\Phi$ such that $\mathcal{A} \cong \mathcal{B}$, contradicting stipulation 4.11.1. $\dashv$