CHAPTER 5

Ehrenfeucht-Fraïssé games

This chapter will be devoted to a game-theoretic characterisation of the notion of elementary equivalence. These interpretations in terms of games are not just fun: they can often be applied in situations where other methods fail.

Throughout this chapter we will, for simplicity, be working in a finite language without function symbols.

1. Definition of the game

DEFINITION 5.1. Given two models M and N and a natural number $n \in \mathbb{N}$ we define a game as follows. It is a two-player game in which two players, player I (who is male) and player II (who is female), move in turn. Player I starts and the game ends after n rounds, so after both players have played n moves. A move by a player consists of picking an element from one of the two structures. Player I has complete freedom and can pick an element from whichever structures he likes, but player II always has to reply by picking an element from the other structure (that is, player II is not allowed to respond by picking an element from the same structure as the one player I just played in). So if in round i player I chooses an element $a_i \in M$, player II replies by picking an element $b_i \in N$, and if in round i player I chooses an element $b_i \in N$, then player II replies by picking an element $a_i \in M$. After n rounds the two players have constructing two sequences $(a_1, \ldots, a_n)$ and $(b_1, \ldots, b_n)$ of elements from M and N, respectively. Player II wins if $\{(a_i, b_i) \colon 1 \leq i \leq n\}$ is a well-defined injective function $f \colon \{a_1, \ldots, a_n\} \to N$ and, moreover, this function is a local isomorphism; otherwise player I wins. We denote this game by $\mathcal{G}_n(M, N)$ and we call it an *Ehrenfeucht-Fraïssé game*.

Let us first remark that:

PROPOSITION 5.2. *One of the players has a winning strategy in* $\mathcal{G}_n(M, N)$.

This is a consequence of a general result in game theory:

THEOREM 5.3. (Zermelo) *In a two-player game of perfect information in which there are no infinite plays and no ties, one of the two players has a winning strategy.*

PROOF. (Sketch) Let us say that a position in the game is *losing* for a player if in that position the other player has a winning strategy. The idea is that if none of the two players has a winning strategy, both can play in such a way that they avoid any losing positions. That is, if player II does not have a winning strategy, then player I can play a move in the initial position of the game after which the position is not lost for him. But if player I also does not have a winning strategy, the position after he has played this move is also not lost for player II. That means that player II can reply by playing a move after which the position is not lost for her. But after player II has played such a move, the position is also not lost for player I:

because otherwise the position just before player II played her move must have been lost for player I. This means that player I can reply by playing a move after which the position is not lost for him; actually, it will not be lost for either of the two players. By proceeding in this vein both players end up playing a game of infinite length, which contradicts the assumption that any possible way of playing the game ends after a finite number of moves in a win for one of the two players. $\square$

The reason we are interested in Ehrenfeucht-Fraïssé games is that they allow us to characterise elementary equivalence.

THEOREM 5.4. *Let L be a finite language without function symbols and let M and N be L-structures. Then $M \equiv N$ if and only if the player II has a winning strategy in $\mathcal{G}_n(M, N)$ for all n.*

A more refined statement is true. To formulate it, we need the following definition.

DEFINITION 5.5. The *quantifier-depth* $\mathrm{dp}(\varphi)$ of a formula φ is defined inductively as follows:

- $\mathrm{dp}(\varphi) = 0$ if φ is atomic.
- $\mathrm{dp}(\varphi \Box \psi) = \max\{\mathrm{dp}(\varphi), \mathrm{dp}(\psi)\}$ for $\Box \in \{\wedge, \vee, \rightarrow\}$.
- $\mathrm{dp}(\neg\varphi) = \mathrm{dp}(\varphi)$,
- $\mathrm{dp}(\exists x\, \varphi) = \mathrm{dp}(\forall x\, \varphi) = \mathrm{dp}(\varphi) + 1$.

We will write $M \equiv_n N$ if

$$M \models \varphi \Leftrightarrow N \models \varphi$$

for any sentence φ with quantifier-depth at most n.

THEOREM 5.6. *Let L be a finite language without function symbols and let M and N be L-structures. Then $M \equiv_n N$ if and only if the player II has a winning strategy in $\mathcal{G}_n(M, N)$.*

The proof of this theorem is a bit finicky: we will give it in Section 3. But before we give this proof, let us first discuss an application.

2. An application

We give one application of Theorem 5.4. Let $L = \{<\}$ and let T be the L-theory asserting that $<$ is a discrete linear order without greatest or smallest element. Discreteness means:

$$\forall x\, \exists y_0\, \exists y_1 \left(y_0 < x < y_1 \wedge \forall z\, (z < x \rightarrow z \leq y_0 \wedge x < z \rightarrow y_1 \leq z\,) \right),$$

where $x \leq y$ abbreviates $x < y \vee x = y$. In other words, discreteness means that each element has an immediate successor and predecessor. For example, $(\mathbb{Z}, <)$ is a model of T.

We claim that T is a complete theory, or, equivalently that any model N of T is elementarily equivalent to $(\mathbb{Z}, <)$.

PROPOSITION 5.7. *The theory T of discrete linear orders with no top or bottom element is a complete theory. In particular, $(\mathbb{Z}, <) \models \varphi$ if and only if $T \models \varphi$ for all L-sentences φ.*

PROOF. We are going to use games. But before we do this, we should first try to understand the general structure of a model N of T.

For elements $a, b \in N$ let us write aEb if b is the nth successor or predecessor of a for some natural number n. Then E is an equivalence relation and each E-class is a linear order that looks like $(\mathbb{Z}, <)$. In addition, the collection of E-classes is linearly ordered as well (by saying that $[a] < [b]$ if $\neg E(a, b) \wedge a < b$). This means that every model of T is of the form $(L \times \mathbb{Z}, <)$, where L is a linear order and $<$ is the lexicographic order on $L \times \mathbb{Z}$ (that is, $(i, a) < (j, b)$ if $i < j$, or both $i = j$ and $a < b$). Conversely, every linear order of this form is a model of T.

So let M be $(\mathbb{Z}, <)$, and let N be $L \times \mathbb{Z}$ with the lexicographic order, where L is any linearly ordered set. We wish to show that $M \equiv N$ and we do this by supplying for each natural number n a winning strategy for player II in the game $\mathcal{G}_n(M, N)$.

If $a, b \in \mathbb{Z}$, we define the distance between a and b to be $\mathrm{dist}(a, b) = |b - a|$, and if $x = (i, a), y = (j, b) \in L \times \mathbb{Z}$, we define the distance to be $\mathrm{dist}(x, y) = |b - a|$ if $i = j$ and $\mathrm{dist}(a, b) = \infty$ if $i \neq j$. The problem for player II is that player I can play elements that are infinitely far apart in N and force player II to play elements that are finitely far apart in M. The crux is that the number of rounds that the game will last has been fixed in advance (and player II knows this number) and if player II can play elements that are sufficiently far apart to avoid conflicts, then she can win the game. Indeed, we claim that if the game lasts i more rounds then Player II only has to ensure that distances $< 2^i$ are preserved. More precisely, player II can win by ensuring the following condition:

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

Clearly, if player II can actually achieve this, she will win because after n rounds there will be a local isomorphism.

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

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

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

3. A proof

The aim of this section give a proof (sketch) for Theorem 5.11. We start off with some syntactic considerations.

A formula is a *boolean combination* of formulas in S if it can be obtained from S by applying conjunction, disjunction, implication and negation (that is, all the possible propositional operations). In addition, let us say that a collection of formulas S is *finite up to logical equivalence* if there is a finite set of formulas $S_0 \subseteq S$ such that each element in S is logically equivalent to some element in S_0.

LEMMA 5.8. *Let A be a collection of formulas and assume B consists of all boolean combinations of A.*

(i) *If M and N make the same formulas in A true, then they also make the same formulas in B true.*
(ii) *If A is finite up to logical equivalence, then so is B.*

PROOF. (i) is proved by induction on the logical complexity of the formulas in B.

(ii): Suppose each element in A is logically equivalent to some element of $\{\varphi_0, \ldots, \varphi_{n-1}\}$. Then each element $\psi \in B$ is equivalent to a formula of the form

$$\bigvee_{\sigma \in S} \Big(\bigwedge_{\{i \,:\, \sigma(i)=1\}} \varphi_i \wedge \bigwedge_{\{i \,:\, \sigma(i)=0\}} \neg\varphi_i \Big)$$

for some $S \subseteq \{0,1\}^n$. Details are left to the reader. □

DEFINITION 5.9. Let us say that for a set of variables $\{x_1, \ldots, x_m\}$ and a natural number n a formula φ is *special* if:

(1) either $n = 0$ and φ is an atomic formula with free variables among $\{x_1, \ldots, x_m\}$, or
(2) $n > 0$ and φ is of the form $\exists x_{m+1}\, \varphi$ where φ is a formula with quantifier-depth at most $n-1$ and free variables among $\{x_1, \ldots, x_{m+1}\}$.

LEMMA 5.10. *Let L be a finite language without function symbols.*

(i) *Every L-formula with quantifier-depth at most n and free variables among $\{x_1, \ldots, x_m\}$ is logically equivalent to a boolean combination of special formulas with respect to $\{x_1, \ldots, x_m\}$ and n.*
(ii) *The collection of L-formulas with quantifier-depth at most n and free variables among $\{x_1, \ldots, x_m\}$ is finite up to logical equivalence.*

PROOF. (i) Each quantifier-free formula is a boolean combination of atomic formulas, and each formula of quantifier-depth at most $n+1$ is a boolean combination of formulas of the form $\exists x\, \varphi$ and $\forall x\, \varphi$ with φ having quantifier-depth at most n. Up to logical equivalemce, we can rename variables so that x becomes x_{m+1} and eliminate $\forall x\, \varphi$ in favour of $\neg\exists x\, \neg\varphi$.

(ii) is proved by induction. For $n = 0$ observe that the number of atomic formulas with free variables among $\{x_1, \ldots, x_m\}$ is finite (here we use that L is a finite language without function symbols). So the desired statement follows from point (i) in this lemma and point (ii) in the previous lemma. For the induction step we argue in the same way, with special formulas instead of atomic formulas. □

THEOREM 5.11. *Let L be a finite language without function symbols and let M and N be L-structures. Then $M \equiv_n N$ if and only if the player II has a winning strategy in $\mathcal{G}_n(M, N)$.*

PROOF. (Sketch) ⇒: Suppose $M \equiv_n N$. We will outline a winning strategy for player II. Suppose player I plays a move $a \in M$ (the case that player I plays an element $b \in N$ is similar). We claim that player II can choose an element $b \in N$ in such a way that $(M, a) \equiv_{n-1} (N, b)$. (This statement should be understood in the following way: if we would add a constant c to the language and interpret it as a in M and b in N, then M and N are $(n-1)$-elementary equivalent in the extended language.)

Indeed, suppose that $\{\varphi_i(x_0) : i \in I\}$ is a finite set of formulas of quantifier-depth at most $n-1$ and with free variable $\{x_0\}$ and suppose that each such formula is equivalent to an element in this set. Put

$$\psi(x_0) := \bigwedge\{\varphi_i(x_0) : M \models \varphi_i(a)\} \wedge \bigwedge\{\neg\varphi_i(x_0) : M \not\models \varphi_i(a)\}.$$

Then $M \models \exists x_0\, \psi(x_0)$, because $M \models \psi(a)$. Since $M \equiv_n N$ and $\exists x_0\, \psi(x_0)$ is of quantifier-depth n, there is an element $b \in N$ such that $N \models \psi(a)$. But then $(M, a) \equiv_{n-1} (N, b)$, as desired.

If player II continues in this way, the players will end up producing sequences $(a_1, \ldots, a_n)$ in M and $(b_1, \ldots, b_n)$ in N such that

$$(M, a_1, \ldots, a_n) \equiv_0 (N, b_1 \ldots, b_n).$$

But then the function $f(a_i) = b_i$ is a local isomorphism, so player II will win the game.

$\Leftarrow$: Suppose $M \not\equiv_n N$. Now we will outline a winning strategy for player I.

If $M \not\equiv_n N$ there must a sentence of the form $\exists x\, \psi(x)$ where ψ has quantifier-depth $n-1$ which is true in one structure, but not in the other. Suppose it is true in M (the case where it is true in N is similar). Then player I starts by picking an element $a \in M$ such that $M \models \psi(a)$. Player II has to respond by picking an element $b \in N$. But then $N \not\models \psi(b)$, so $(M, a) \not\equiv_{n-1} (N, b)$.

If player I continues in this way, the players will end up producing sequences $(a_1, \ldots, a_n)$ in M and $(b_1, \ldots, b_n)$ in N such that

$$(M, a_1, \ldots, a_n) \not\equiv_0 (N, b_1 \ldots, b_n).$$

But then $f(a_i) = b_i$, even when this defines a function, cannot be a local isomorphism. Therefore player I wins the game. $\square$

4. Exercises

Exercise 1. Give a direct proof of Proposition 5.2, that is, without using Theorem 5.3. *Hint:* Simply use induction on n.

Exercise 2. Let L be the first-order language of linear ordering. Show that if $h < 2^k$ then there is a formula $\varphi(x, y)$ of L of quantifier depth $\leq k$ which expresses (in any linear ordering) "$x < y$ and there are at least h elements strictly between x and y".

Exercise 3. The circle of length $N \in \mathbb{N}$ is the structure $\mathcal{C}_N := (C_N, R)$, where $C_N = \{0, \ldots, N-1\}$ and $R = \{(i, j) \in C_N \times C_N : j = i + 1 \mod N\}$.

(a) Give a function $f : \mathbb{N} \to \mathbb{N}$ such that $\mathcal{C}_N \equiv_n \mathcal{C}_{N'}$ whenever $N, N' \geq f(n)$.
(b) Is there a first-order formula φ such that $\mathcal{C}_N \models \varphi$ if and only if N is even?

Exercise 4. Show that there is no formula of first-order logic which expresses "(a, b) is in the transitive closure of R", even on finite structures. (For infinite structures it is easy to show there is no such formula.)