Ehrenfeucht-Fraïssé games

Sam van Gool

May 13, 2026

These are lecture notes for the part of the L3 course *Logique* at ENS Paris-Saclay in which we talk about Ehrenfeucht-Fraïssé games. The main result is the characterization of elementary equivalence in terms of these games(Theorem 27). We will also see applications to definability. The exposition is strongly inspired by [2, Ch. 1,2][1], also see [1, Ch. XII].

Notation 1. Throughout these notes, we consider first-order logic over an arbitrary *finite* and *relational* signature $\mathcal{L}$. The assumption that $\mathcal{L}$ does not contain function symbols simplifies some technical points, and the general case can be reduced to this one. The assumption that $\mathcal{L}$ is finite, however, will be used in an essential wayin Lemma 25. Since $\mathcal{L}$ is fixed, we will write 'structure' instead of '$\mathcal{L}$-structure'. When $\mathcal{A}$, $\mathcal{B}$, ... are structures, we denote their underlying sets by A, B, ...

1 Local isomorphisms

Our goal is to understand elementary equivalence in a semantic way. We begin with a notion that is strictly stronger than elementary equivalence, namely, that of an isomorphism.

Definition 2. Let $\mathcal{A}, \mathcal{B}$ be structures. An *isomorphism* from $\mathcal{A}$ to $\mathcal{B}$ is a bijective function $f\colon A \to B$ such that, for any n-ary relation symbol R and any $(a_1, \ldots, a_n) \in A^n$, we have

$$R^{\mathcal{A}}(a_1, \ldots, a_n) \text{ if, and only if, } R^{\mathcal{B}}(fa_1, \ldots, fa_n) \ .$$

Definition 3. Let $\mathcal{A}$ be a structure and let $A' \subseteq A$. The *restriction* of $\mathcal{A}$ to A' is the structure $\mathcal{A}'$ with underlying set A' and, for each n-ary relation symbol R,

$$R^{\mathcal{A}'} := R^{\mathcal{A}} \cap (A')^n \ .$$

Definition 4. We define by induction a notion of *p-isomorphism*, for every natural number $p \geq 0$:

- A 0-*isomorphism* from $\mathcal{A}$ to $\mathcal{B}$ is an isomorphism from a finite restriction of $\mathcal{A}$ to a finite restriction of $\mathcal{B}$.
- A $(p+1)$-*isomorphism* from $\mathcal{A}$ to $\mathcal{B}$ is a 0-isomorphism f from $\mathcal{A}$ to $\mathcal{B}$ such that:
  - (*forth*) for any $a \in A$, there exists an extension g of f such that $a \in \mathrm{dom}(g)$, and g is a p-isomorphism;

[1] The book [2] is also available in English [3], with a long new introduction and an additional new section.

– (*back*) for any $b \in B$, there exists an extension g of f such that $b \in \mathrm{im}(g)$, and g is a p-isomorphism.

An alternative name for '0-isomorphism' is *local isomorphism*.

Notation 5. We write $f\colon \mathcal{A} \cong \mathcal{B}$ if f is an isomorphism and $f\colon \mathcal{A} \cong_p \mathcal{B}$ if f is a p-isomorphism.

Remark 6. For any isomorphism $f\colon \mathcal{A} \cong \mathcal{B}$, and for any finite subset A' of A, the restriction of f to A' is a p-isomorphism for every $p \geq 0$; the proof is by induction on p.

Remark 7. For any p-isomorphism $f\colon \mathcal{A} \cong_p \mathcal{B}$, the restriction of f to a subset A' of $\mathrm{dom}(f)$ is also a p-isomorphism; the proof is by induction on p. In particular, if there exists any p-isomorphism from $\mathcal{A}$ to $\mathcal{B}$, then the empty function is also a p-isomorphism from $\mathcal{A}$ to $\mathcal{B}$.

Definition 8. For $n \in \mathbb{N}$, an *n-pointed structure* is a pair $\mathcal{A}, \overline{a}$, where $\mathcal{A}$ is a structure and $\overline{a} \in A^n$.

Notation 9. When $\overline{a}$ and $\overline{b}$ are tuples of the same length in structures $\mathcal{A}$ and $\mathcal{B}$, respectively, we write $\overline{a} \mapsto \overline{b}$ for the relation $\{(a_i, b_i) \mid 1 \leq i \leq n\} \subseteq A \times B$.

Definition 10. Let $\mathcal{A}, \overline{a}$ and $\mathcal{B}, \overline{b}$ be n-pointed structures and $p \geq 0$. Then $\mathcal{A}, \overline{a}$ is *p-equivalent* to $\mathcal{B}, \overline{b}$, denoted $\mathcal{A}, \overline{a} \sim_p \mathcal{B}, \overline{b}$, provided the relation $\overline{a} \mapsto \overline{b}$ is a p-isomorphism.

Remark 11. The relation $\sim_p$ is an equivalence relation: an exercise in manipulating Definition 4.

Notation 12. In Definition 10, when $\mathcal{A}$ and $\mathcal{B}$ are clear from the context, we also just write $\overline{a} \sim_p \overline{b}$. In case $n = 0$, so that $\overline{a}$ and $\overline{b}$ are empty tuples, we write $\mathcal{A} \sim_p \mathcal{B}$.

Remark 13. Note that, if $\overline{a} \mapsto \overline{b}$ is a p-isomorphism, then $\overline{a} \mapsto \overline{b}$ is in particular a bijective function, so, for any $1 \leq i, j \leq n$, we have $a_i = a_j$ if, and only if, $b_i = b_j$.

Example 14. Let $\mathcal{A}$ be a finite structure, with $p := |A|$. We show that, if $\mathcal{A}$ is $(p+1)$-equivalent to $\mathcal{B}$, then $\mathcal{A}$ is isomorphic to $\mathcal{B}$. Apply the forth condition p times to obtain a 1-isomorphism f with domain equal to A. In particular, since f is a 0-isomorphism, f must be injective. We claim that f is also surjective. Towards a contradiction, suppose that there were $b \in B \setminus \mathrm{im}(f)$. Since f is a 1-isomorphism, we can pick an extension g of f which is a 0-isomorphism and has $b \in \mathrm{im}(g)$. Pick $a \in A$ such that $g(a) = b$. But since $\mathrm{dom}(f) = A$, we also have $f(a) = b'$, so $b' \in \mathrm{im}(f)$, and $b' \neq b$, which is impossible since g extends f. Thus, g is a bijection from A to B (in particular, B is also finite) and a 0-isomorphism, and therefore an isomorphism from $\mathcal{A}$ to $\mathcal{B}$.

Definition 15. Let $\mathcal{A}, \overline{a}$ and $\mathcal{B}, \overline{b}$ be n-pointed structures. We say that $\mathcal{A}, \overline{a}$ and $\mathcal{B}, \overline{b}$ are *ω-equivalent* if, for all $p \geq 0$, we have $\mathcal{A}, \overline{a} \sim_p \mathcal{B}, \overline{b}$.

2 Games

The definitions in Section 1 are due to Fraïssé. We now reformulate them equivalently in a game-theoretic way; this reformulation is due to Ehrenfeucht.

Definition 16. Let $n \in \mathbb{N}$ and let $\mathcal{A}, \overline{a}$ and $\mathcal{B}, \overline{b}$ be n-pointed structures. For $p \geq 0$, the *p-round Ehrenfeucht-Fraïssé (EF) game on $\mathcal{A}, \overline{a}$ and $\mathcal{B}, \overline{b}$* is a game with two players. In each round, player 1 chooses an element of one of the two structures, and player 2 responds by choosing an element of the other structure. After p rounds, we obtain two p-tuples, $(c_1, \ldots, c_p) \in A^p$ and $(d_1, \ldots, d_p) \in B^p$. By definition, player 2 wins this play of the game if, and only if, the relation $\{(a_i, b_i) \mid 1 \leq i \leq n\} \cup \{(c_i, d_i) \mid 1 \leq i \leq p\}$ is a 0-isomorphism.

Proposition 17. *For any $p, n \geq 0$ and any n-pointed structures $\mathcal{A}, \overline{a}$ and $\mathcal{B}, \overline{b}$, we have $\mathcal{A}, \overline{a} \sim_p \mathcal{B}, \overline{b}$ if, and only if, player 2 has a winning strategy in the p-round EF game on $\mathcal{A}, \overline{a}$ and $\mathcal{B}, \overline{b}$.*

Proof. By induction on p. If $p = 0$, then the game ends without any action from the players, and the statement follows from the definitions. Now assume the statement holds for $p \geq 0$; we show the equivalence for $p + 1$.

Suppose that $\overline{a} \mapsto \overline{b}$ is a $(p + 1)$-isomorphism. We describe a strategy for player 2. In the first round of the game, there are two cases: player 1 chooses an element of A, or he chooses an element of B. We only treat the first case, the second is similar. Suppose that player 1 chooses $c_1 \in A$. By the forth condition, pick an extension g of $\overline{a} \mapsto \overline{b}$ whose domain contains c_1 and such that g is a p-isomorphism. Player 2 will play $d_1 := g(c_1)$. Since g is a p-isomorphism, using Remark 7, its restriction $\overline{a} \cdot c_1 \mapsto \overline{b} \cdot d_1$ is also a p-isomorphism. Thus, $\mathcal{A}, \overline{a} \cdot c_1 \sim_p \mathcal{B}, \overline{b} \cdot d_1$. By the induction hypothesis, player 2 has a winning strategy for the p-round EF game on $\mathcal{A}, \overline{a}$ and $\mathcal{B}, \overline{b}$, which she will now follow.

Conversely, suppose that player 2 has a winning strategy in the $(p + 1)$-round game on $\mathcal{A}, \overline{a}$ and $\mathcal{B}, \overline{b}$. We need to show that $\overline{a} \mapsto \overline{b}$ is a $(p + 1)$-isomorphism. First, since player 2 can win the game, $\overline{a} \mapsto \overline{b}$ must in particular be a 0-isomorphism, being the restriction of a 0-isomorphism (Remark 7). We now verify the forth condition; the back condition is similar. Let $c_1 \in A$. Pick $d_1 \in B$ according to the winning strategy of player 2, when player 1 plays c_1 in the first round. Then player 2 still has a winning strategy in the p-round game on $\mathcal{A}, \overline{a} \cdot c_1$ and $\mathcal{B}, \overline{b} \cdot d_1$. By the induction hypothesis, $\overline{a} \cdot c_1 \mapsto \overline{b} \cdot d_1$ is thus a p-isomorphism. □

3 Linear orders

As useful examples for getting an intuition about the above definitions, we will now examine the games on linear orders.

Definition 18. A *linear order* is a structure $(A, \leq)$ such that $\leq$ is reflexive, transitive, anti-symmetric, and total.

A linear order $(A, \leq)$ is *dense* if for all $a, b \in A$, if $a < b$ then there exists $c \in A$ such that $a < c < b$.

A *minimum* is an element $a \in A$ such that, for all $b \in A$, $a \leq b$, and a *maximum* is defined dually.

A linear order is *without endpoints* if it has neither a minimum, nor a maximum.

Example 19. The linear order of rational numbers $\mathcal{Q} := (\mathbb{Q}, \leq^{\mathbb{Q}})$ is a dense linear order without endpoints.

Definition 20. When a, b are elements of a linear order, we say that b is an (immediate) *successor* of a provided that $a < b$ and there does not exist $c \in A$ such that $a < c < b$. In this case, a is called an (immediate) *predecessor* of b.

A linear order $(A, \leq)$ is *discrete* if any non-maximum a has a successor, and any non-minimum a has a predecessor.

Example 21. The linear order of integers $\mathcal{Z} := (\mathbb{Z}, \leq^{\mathbb{Z}})$ is a discrete linear order without endpoints.

We observe that $\mathcal{Z}$ is 2-equivalent to $\mathcal{Q}$, but not 3-equivalent: player 1 can first play, for instance, 0 and 1 in $\mathcal{Z}$, to which player 2 responds with some $b_1, b_2 \in \mathbb{Q}$ with $b_1 < b_2$. In the third round, player 1 plays some $b_3 \in \mathbb{Q}$ with $b_1 < b_3 < b_2$, and any response by player 2 will be losing.

Proposition 22. *Let $\mathcal{A} = (A, \leq^A)$ be any structure in the signature with one binary relation. The following are equivalent:*

1. $\mathcal{A} \sim_3 \mathcal{Q}$;
2. $\leq^A$ *is a dense linear order without endpoints;*
3. $\mathcal{A} \sim_\omega \mathcal{Q}$.

In particular, any two dense linear orders without endpoints are p-equivalent, for all $p \geq 0$.

Proof. (1) $\Rightarrow$ (2). We first show $\leq^A$ is linear. let $a, b \in A$ be arbitrary. Following the winning strategy of player 2 in the 3-round game on $\mathcal{A}$ and $\mathcal{Q}$, pick $q, r \in \mathbb{Q}$ such that $ab \mapsto qr$ is a 0-isomorphism. In particular, since either $q \leq r$ or $r \leq q$, the same holds for a and b, so $\leq^A$ is linear.
The proofs of reflexivity, transitivity, and antisymmetry are similar.
To see that $\leq^A$ does not have a minimum element, let $a \in A$ be arbitrary. Player 1 first plays a, say that player 2 responds q. Now player 1 plays $q - 1$. Player 2 must respond with an element strictly below a. The proof that $\leq^A$ does not have a maximum is analogous.
Finally, to see that $\leq^A$ is dense, let $a, b \in A$ be arbitrary such that $a < b$. Player 1 first plays a and b, and player 2 responds with q, r, respectively. We must have $q < r$. Pick $s \in \mathbb{Q}$ such that $q < s < r$, and let player 1 play s in the third and final round. Player 2 responds with some $c \in A$ such that $a < c < b$.
(2) $\Rightarrow$ (3). We show by induction on p that player 2 can always maintain the invariant which says that, for the tuples $(a_1, \dots, a_p)$ and $(q_1, \dots, q_p)$ that have been played so far, $a_i \leq a_j$ if, and only if, $q_i \leq q_j$, for all $1 \leq i, j \leq p$. The base case is trivial. Assume that player 2 has succeeded for p rounds, and now player 1 plays $q \in \mathbb{Q}$, say; the case were player 1 plays $a \in A$ is symmetric. If there exists $1 \leq i \leq p$ such that $q = q_i$, then player 2 plays a_i. Suppose now that $q \neq q_i$ for all $1 \leq i \leq p$. Using the linearity of $\mathbb{Q}$, we must be in one of the following three cases:

- $q < q_i$ for all $1 \leq i \leq p$;
- $q > q_i$ for all $1 \leq i \leq p$;
- there exist $1 \leq i, j \leq p$ such that $q_i < q < q_j$ and there is no $1 \leq k \leq p$ such that $q_i < q_k < q_j$.

In the first case, player 2 plays some element below $\min\{a_i \mid 1 \leq i \leq p\}$, using that $\leq^A$ does not have a minimum, and in the second case, player 2 plays some element above $\max\{q_i \mid 1 \leq i \leq p\}$, using that $\leq^A$ does not have a maximum. In the third case, the invariant implies $a_i < a_j$, so player 2 can choose an element a strictly between a_i and a_j, since $\leq^A$ is dense.
(3) $\Rightarrow$ (1) is trivial. $\square$

4 Fraïssé's theorem

Definition 23. The *quantifier rank* of a first-order formula φ is the maximal nesting depth of quantifiers in φ. More formally, qr is a function from formulas to natural numbers, defined inductively, for φ and ψ formulas, by:

- if φ is atomic, $\mathsf{qr}(\varphi) := 0$,

- $\mathsf{qr}(\neg\varphi) := \mathsf{qr}(\varphi)$,
- $\mathsf{qr}(\varphi \vee \psi) := \max\{\mathsf{qr}(\varphi), \mathsf{qr}(\psi)\}$,
- $\mathsf{qr}(\exists x\varphi) := \mathsf{qr}(\varphi) + 1$.

Remark 24. Let $p \geq 0$. If $\mathsf{qr}(\varphi) \leq p+1$, then φ is a Boolean combination of formulas of the form $\exists x\psi$, where $\mathsf{qr}(\psi) \leq p$.

Lemma 25. *Suppose that $\mathcal{L}$ is a finite relational signature. For any $n, p \geq 0$, the equivalence relation $\sim_p$ on n-pointed structures has finitely many equivalence classes.*

Proof. By induction on p. The 0-class of an n-pointed structure $\mathcal{A}, \overline{a}$ is entirely determined by the restriction of the relations to the tuple $\overline{a}$. There is a finite number of structures on an n-tuple of (not necessarily distinct) elements, so the number of 0-classes is finite. The $(p+1)$-class of $\mathcal{A}, \overline{a}$ is entirely determined by the set of all p-equivalence classes that can be obtained by extending $\overline{a}$ with one more element; by the induction hypothesis, there are finitely many p-equivalence classes of $(n+1)$-pointed structures, and we thus obtain the result. □

Remark 26. Suppose that the arities of the relations in $\mathcal{L}$ are $r_1, \ldots, r_k$. We can compute bounds on the number of equivalence classes of $\sim_p$. For $n, p \geq 0$, let $C(n,p)$ be the cardinality of the set of p-classes of n-pointed structures. There exist $2^{m^{r_1}} \times \cdots \times 2^{m^{r_k}}$ structures on a set with m elements. On a given n-tuple, equality between its elements determines an equivalence relation on the set of n indices. We get $C(n,0) = \sum_{m=1}^{n} S_{n,m} 2^{m^{r_1}} \times \cdots \times 2^{m^{r_k}}$ distinct structures on an n-tuple, where $S_{n,m}$ denotes the number of equivalence relations on an n-element set having precisely m classes ($S_{n,m}$ is called the *Stirling number of the second kind*). The proof of Lemma 25 moreover shows that $C(n, p+1) \leq 2^{C(n+1,p)}$.

Theorem 27. *For any $p \geq 0$, for any $n \geq 0$, and any n-pointed structures $\mathcal{A}, \overline{a}$ and $\mathcal{B}, \overline{b}$, the following are equivalent:*

1. *$\mathcal{A}, \overline{a} \sim_p \mathcal{B}, \overline{b}$;*
2. *$\mathcal{A}, \overline{a}$ and $\mathcal{B}, \overline{b}$ satisfy the same formulas with n free variables of quantifier rank $\leq p$.*

Proof. (1) ⇒ (2). By induction on p. When $p = 0$, first note that $\overline{a}$ and $\overline{b}$ satisfy the same atomic formulas since $\overline{a} \mapsto \overline{b}$ is a 0-isomorphism. Thus, $\overline{a}$ and $\overline{b}$ satisfy the same quantifier-free formuals, since these are Boolean combinations of atomic formulas. For the inductive case, suppose the implication has been proved for p; we prove it for $p+1$. By Remark 24, it suffices to prove that, for any ψ with $\mathsf{qr}(\psi) \leq p$, we have $\mathcal{A}, \overline{a} \models \exists x\psi$ if, and only if, $\mathcal{B}, \overline{b} \models \exists x\psi$. Let $\psi(\overline{y}, x)$ be a formula of quantifier rank $\leq p$, and suppose that $\mathcal{A}, \overline{a} \models \exists x\psi(\overline{y}, x)$. Pick $a_{p+1} \in A$ such that $\mathcal{A}, \overline{a}a_{p+1} \models \psi(\overline{y}, x)$. Since $\mathcal{A}, \overline{a} \sim_{p+1} \mathcal{B}, \overline{b}$, using the forth condition, pick $b_{p+1} \in B$ such that $\mathcal{A}, \overline{a}a_{p+1} \sim_p \mathcal{B}, \overline{b}b_{p+1}$. By the induction hypothesis, $\mathcal{B}, \overline{b}b_{p+1} \models \psi(\overline{y}, x)$. Thus, $\mathcal{B}, \overline{b} \models \exists x\psi(\overline{y}, x)$. The other direction is similar, using the back condition.
(2) ⇒ (1). We will show by induction on p that, for any p-equivalence class C of n-pointed structures, there is a formula $\varphi_C(x_1, \ldots, x_n)$ with $\mathsf{qr}(\varphi) \leq p$ such that the n-pointed structures in C are exactly those satisfying φ_C. The implication immediately follows from this.
For the case $p = 0$, let $\mathcal{A}, \overline{a}$ be an n-pointed structure. Write $\Phi_{n,0}$ for the set of literals (i.e., atomic or negated atomic formulas) in variables $x_1, \ldots, x_n$; note that $\Phi_{n,0}$ is finite. Let $\varphi_{[\mathcal{A},\overline{a}]_0}$ be

the conjunction of all the formulas in $\Phi_{n,0}$ that are satisfied by $\mathcal{A}, \overline{a}$:

$$\varphi_{[\mathcal{A},\overline{a}]_0} := \bigwedge\{\varphi \in \Phi_{n,0} \mid \mathcal{A}, \overline{a} \models \varphi\} \ .$$

Now assume the statement for p; we prove it for $p+1$. Let $n \geq 0$ and let $\mathcal{A}, \overline{a}$ be an n-pointed structure. Consider the set of p-equivalence classes $\mathcal{C} := \{[\mathcal{A}, \overline{a}a_{n+1}]_p \mid a_{n+1} \in A\}$, which is finite by Lemma 25. By the induction hypothesis, for each $C \in \mathcal{C}$, pick a formula φ_C of quantifier rank $\leq p$ such that the $(n+1)$-pointed structures in C are exactly those satisfying φ_C. Define the set of formulas $\Phi_{n,p+1} := \{\exists x_{n+1}\varphi_C \mid C \in \mathcal{C}\} \cup \{\neg\exists x_{n+1}\varphi_C \mid C \in \mathcal{C}\}$, and let

$$\varphi_{[\mathcal{A},\overline{a}]_p} := \bigwedge\{\varphi \mid \varphi \in \Phi_{n,p+1} \text{ and } \mathcal{A}, \overline{a} \models \varphi\} \ .$$

This formula is clearly of quantifier rank $\leq p+1$, and it is satisfied by $\mathcal{B}, \overline{b}$ if, and only if, $\mathcal{A}, \overline{a} \sim_p \mathcal{B}, \overline{b}$. □

Remark 28. Direction (2) ⇒ (1) produces a characteristic formula for each p-equivalence class C of n-pointed structures, which is a conjunction of formulas in the set $\Phi_{n,p}$. It follows that any formula φ of quantifier rank $\leq p$ is equivalent to a disjunction of conjunctions of formulas in $\Phi_{n,p}$. Indeed, for any formula $\varphi(x_1, \ldots, x_n)$ with $\mathsf{qr}(\varphi) \leq p$, let

$$[\![\varphi]\!] := \{\mathcal{A}, \overline{a} \mid \mathcal{A}, \overline{a} \models \varphi\} \ ,$$

and consider

$$\varphi' := \bigvee_{\mathcal{A},\overline{a} \in [\![\varphi]\!]} \varphi_{[\mathcal{A},\overline{a}]_p} \ .$$

We claim that φ' is equivalent to φ. First, if $\mathcal{B}, \overline{b} \models \varphi'$, then $\mathcal{B}, \overline{b} \models \varphi_{[\mathcal{A},\overline{a}]_p}$ for some $\mathcal{A}, \overline{a} \in [\![\varphi]\!]$, so that $\mathcal{B}, \overline{b} \sim_p \mathcal{A}, \overline{a}$, by the proof of (2) ⇒ (1) of Theorem 27. Since φ is of quantifier rank $\leq p$ and $\mathcal{A}, \overline{a} \models \varphi$, it now follows that $\mathcal{B}, \overline{b} \models \varphi$ from (1) ⇒ (2) in Theorem 27. Conversely, if $\mathcal{B}, \overline{b} \models \varphi$, then $\mathcal{B}, \overline{b} \in [\![\varphi]\!]$, and clearly $\mathcal{B}, \overline{b} \models \varphi_{[\mathcal{B},\overline{b}]_p}$ because $\mathcal{B}, \overline{b}$ belongs to its own p-class. Thus, $\mathcal{B}, \overline{b} \models \varphi'$.

References

[1] H-D. Ebbinghaus, J. Flum, and W. Thomas. *Mathematical Logic, Third edition*. Springer, 2021.

[2] B. Poizat. *Cours de théorie des modèles*. 1985.

[3] B. Poizat. *A Course in Model Theory*. Springer, 2000.