Elementary Equivalence, Partial Isomorphisms, and Scott-Karp analysis

These are self-study notes I prepared when I was trying to understand the subject.

1 Elementary equivalence and Finite back and forth

We begin by analyzing the power of finitary analysis of countable structures. In doing so, we introduce finitary versions of all the tools we will see later. This helps to clarify the need for the more powerful infinitary tools.

Definition 1.1 *Let $\mathfrak{A}$ and $\mathfrak{B}$ be two σ-structures. A* partial embedding *between two structures $\mathfrak{A}$ and $\mathfrak{A}$ is a map p whose domain is a finite subset of A and range is a finite subset of B that preserves constants and relations. Namely, p has the following properties:*

(i) *It is injective,*

(ii) *For every constant $c \in \sigma$, $c^{\mathfrak{A}} \in \mathrm{dom}(p)$ and $p(c^{\mathfrak{A}}) = c^{\mathfrak{B}}$, and*

(iii) *For every n-ary relation $R \in \sigma$, and all $a_1, \ldots, a_n \in \mathrm{dom}(p)$,*

$$R^{\mathfrak{A}}(a_1, \ldots, a_n) \Leftrightarrow R^{\mathfrak{B}}(p(a_1), \ldots, p(a_n)).$$

The above definition is purely structural. We could define a partial embedding using language as follows: A partial embedding between two σ-structures $\mathfrak{A}$ and $\mathfrak{B}$ is a finite relation

$$\{(a_1, b_1), \cdots, (a_n, b_n)\} \cup \{(c^{\mathfrak{A}}, c^{\mathfrak{B}}) \colon c \text{ a constant in } \sigma\} \subset A \times B$$

such that the simple extensions $(\mathfrak{A}, a_1, \ldots, a_n)$ and $(\mathfrak{B}, b_1, \ldots, b_n)$ satisfy the same atomic sentences. Equivalently, for all φ atomic, $\mathfrak{A} \models \varphi(a_1, \ldots, a_n) \Leftrightarrow \mathfrak{B} \models \varphi(b_1, \ldots, b_n)$. This definition allows us to define an extension of an embedding p as an embedding q whose graph contains the graph of p.

In general, a partial embedding does not preserve the validity of formulae with quantifiers. However, it does so when it can be "extended" appropriately. The precise meaning of extended is provided by the back and forth property.

Definition 1.2 *A sequence of sets $I_0 \supseteq I_1 \supseteq \cdots \supseteq I_m$ of partial embeddings from $\mathfrak{A}$ to $\mathfrak{B}$ is said to be a m-*Back and Forth system *if for every $j < m$, the following two properties hold:*

(i) *For every $p \in I_{j+1}$ and $a \in A$, there is a $q \in I_j$ that extends p and has a in its domain (Forth property).*

(ii) *For every $p \in I_{j+1}$ and $b \in B$, there is a $q \in I_j$ that extends p and has b in its range (Back property).*

Informally, the above says that a partial embedding in I_m can be extended m times, with the resulting partial embeddings lying in $I_{m-1}, \ldots, I_0$, respectively.

The notion of isomorphism between structures (and therefore of embeddings) is purely structural. It does not *apriori* involve the first order language. Fraisse related this structural notion with sentences in the first order language using the back-and-forth construction above.

Theorem 1.3 (Fraisse) *The following are equivalent:*

(i) *There exists a m-Back and forth system between $\mathfrak{A}$ and $\mathfrak{B}$.*

(ii) *$\mathfrak{A}$ and $\mathfrak{B}$ satisfy the same first order sentences of quantifier rank at most m.*

1.1 Ehrenfeucht Games

Fraisse's formulation of equivalence under first order sentences of a bounded quantifier rank was algebraic. Ehrenfeucht provided a intuitive game-theoretic interpretation. Namely, he showed that two structures satisfy the same first-order sentences of quantifier rank at most m if they are indistinguishable within m rounds of a certain game, known now as the *Ehrenfeucht-Fraisse* game.

The Ehrenfeucht-Fraisse game is played between two players, traditionally called $\forall$ and $\exists$ (or $\forall$belard and $\exists$loise), on two structures $\mathfrak{A}$ and $\mathfrak{B}$. Informally, $\forall$ tries to show that the two structures are different, while $\exists$ aims to show that they are the same (or isomorphic). Player $\forall$ wins if he manages to demonstrate a difference in the two structures before the game ends. If he is unable to do this, $\exists$ wins the game.

The game is played as follows. A positive integer n is fixed ahead of the game. The game consists of an ordered sequence of n rounds, each of which is played as follows: First SPOILER picks an element from either $\mathfrak{A}$ or $\mathfrak{B}$. DUPLICATOR then picks an element from the other structure. The element of $\mathfrak{A}$ picked at the i^{th} round is denoted a_i, and the other element is denoted b_i. Thus, at the end of the game with n rounds, a mapping $a_i \mapsto b_i$ for $1 \leq i \leq n$ has been constructed. The pair $\mathbf{a} = (a_1, \ldots, a_n)$ and $\mathbf{b} = (b_1, \ldots, b_n)$ is called the *play* of the game. Such a n-round Ehrenfeucht-Fraisse game is denoted $\mathrm{EF}_n(\mathfrak{A}, \mathfrak{B})$.

The play $(\mathbf{a}, \mathbf{b})$ is considered a win for the player $\exists$ if the map $p_j \colon \langle \mathbf{a}_{1:j} \rangle_{\mathfrak{A}} \to \langle \mathbf{b}_{1:j} \rangle_{\mathfrak{B}}$ sending a_i to b_i for $1 \leq i \leq j$ and sending each constant $c^{\mathfrak{A}}$ to $c^{\mathfrak{B}}$, is a partial embedding for each j.

The connection between these games and equivalence under formulae of quantifier depth was provided by Ehrenfeucht. To see this connection more clearly, we introduce formulae that encode the m-isomorphism type of a n-tuple $\mathbf{a}$ in $\mathfrak{A}$.

Definition 1.4 *Let $\mathbf{a}$ be an n-tuple from A. Let $\mathbf{x}$ be the sequence of variables $x_1, \ldots, x_n$. Define the m-isomorphism type of $\mathbf{a}$ in $\mathfrak{A}$ by the formulae*

$$\varphi^0_{\mathbf{a}} \triangleq \bigwedge \{\varphi(\mathbf{x}) \colon \varphi \textit{ atomic or negated atomic and } \mathfrak{A} \models \varphi[\mathbf{a}]\},$$

and for $m > 0$,

$$\varphi^m_{\mathbf{a}} \triangleq \bigwedge \{\exists x_{n+1} \varphi^{m-1}_{\mathbf{a}a_{n+1}}(\mathbf{x}, x_{n+1}) \colon a_{n+1} \in A\} \\ \wedge \forall x_{n+1} \{\bigvee \varphi^{m-1}_{\mathbf{a}a_{n+1}}(\mathbf{x}, x_{n+1}) \colon a_{n+1} \in A\}. \quad (1.1)$$

The formulae $\{\varphi_{\mathbf{a}}^m : m \geq 0\}$ encode the isomorphism types of the extensions of the substructure generated by $\mathbf{a}$ in $\mathfrak{A}$. More precisely, $\varphi_{\mathbf{a}}^0$ encodes the isomorphism type of the substructure generated by $\mathbf{a}$ in $\mathfrak{A}$. For $m > 0$, $\varphi_{\mathbf{a}}^m$ describes the isomorphism types the tuple $\mathbf{a}$ can be extended to in m-steps by adding one element of A to the tuple at each step.

The application of these formulae in the game theoretic setting is given by the following lemma.

Lemma 1.5 *Let the notation and setting be as in Defn. 1.4. Let $\mathfrak{B}$ be another σ-structure. Then the following are equivalent.*

(i) DUPLICATOR *wins the m-round Ehrenfeucht-Fraisse game* $\mathrm{EF}_m(\mathfrak{A}, \mathbf{a}, \mathfrak{B}, \mathbf{b})$.

(ii) $\mathfrak{B} \models \varphi_{\mathbf{a}}^m[\mathbf{b}]$.

In particular, $\mathfrak{B} \models \varphi_{\mathbf{a}}^m[\mathbf{b}]$ *if and only if the induced mapping* $\langle \mathbf{a} \rangle_{\mathfrak{A}} \to \langle \mathbf{b} \rangle_{\mathfrak{B}}$ *is a partial embedding.*

The m-isomorphism types of tuples provide a convenient way to talk about the totality of extensions that can be obtained starting with the tuple. They are often used in proofs of Ehrenfeucht's theorem, which we recall next.

Theorem 1.6 (Ehrenfeucht) *The following are equivalent:*

(i) *The player $\exists$ has a winning strategy for the m-round Ehrenfeucht-Fraisse game* $\mathrm{EF}_m(\mathfrak{A}, \mathbf{a}, \mathfrak{B}, \mathbf{b})$.

(ii) *The structures $(\mathfrak{A}, \mathbf{a})$ and $(\mathfrak{B}, \mathbf{b})$ satisfy the same sentences of quantifier depth at most m.*

Note that we get the theorem analogous to Thm. 1.3 by letting both $\mathbf{a}$ and $\mathbf{b}$ be the empty sequences.

The theorems of Fraisse and Ehrenfeucht provide two ways of capturing equivalence under formulae up to a specified quantifier depth. Given one, we can construct the other: namely, given a winning strategy for the m-round game, we can construct a m-back and forth system of partial embeddings, and vice versa. The information in both constructions is the same, just encoded in an algebraic versus a game theoretic setting.

What can we infer if player $\exists$ wins each finite game?

Theorem 1.7 *The following are equivalent:*

(i) *Structures* $\mathfrak{A}$ *and* $\mathfrak{B}$ *are elementarily equivalent.*

(ii) *The player* $\exists$ *has a winning strategy for each of the games* $\mathrm{EF}_m(\mathfrak{A}, \mathfrak{B})$ *for* $m = 0, 1, 2, \ldots$.

At this point, we pause to take stock of the situation. We make two observations.

(i) Having winning strategies for each of the finite games only takes us as far as elementary equivalence. If we wish to go beyond elementary equivalence - to a equivalence that is finer and closer to isomorphism - then we need a stronger criterion than this.

(ii) First order logic is too weak to provide the separation we need for infinite structures. Since equivalence under first order logic can be obtained with a relatively weak constraint - namely, having winning strategies for each finite game.

These observations lead us to consider a logic that is more expressive than first order logic, and consider various means to express equivalence in that logic. We pursue this investigation in the next section.

2 Scott analysis

Next, we introduce Scott formulae, which describe a tuple in a structure up to back-and-forth equivalence till any specified ordinal. Our exposition is based on [Hod93, Bar73, Mar10, Ros82].

Definition 2.1 *Let* $\mathbf{a} = \{a_1, \ldots, a_n\}$ *be a tuple from* A*, and* α *be an ordinal. We define the* Scott Formula $\varphi^{\mathfrak{A}}_{a,\alpha}$*, proceeding by induction on* α *as follows:*

- *For the ordinal* 0*,* $\varphi^{\mathfrak{A}}_{a,0}$ *is defined as the conjunction of all atomic and negated atomic formulae with free variables among* $x_1, \ldots, x_n$ *that are true of* $\mathbf{a}$ *in* $\mathfrak{A}$*.*
- *For successor ordinals* $\alpha = \beta + 1$*, the* Scott Formula $\varphi^{\mathfrak{A}}_{a,\beta}$ *is defined as the conjunction of*
  - (i) $\varphi^{\mathfrak{A}}_{\mathbf{a},\beta}$,
  - (ii) $\bigwedge\{\exists x_{n+1} \varphi^{\mathfrak{A}}_{\mathbf{a}c,\beta}(x_1, \ldots, x_n, x_{n+1})\colon c \in A\}$,
  - (iii) $\forall x_{n+1} \bigvee\{\varphi^{\mathfrak{A}}_{\mathbf{a}d,\beta}(x_1, \ldots, x_n, x_{n+1})\colon d \in A\}$.
- *Finally, for limit ordinals* γ*, we define the Scott formula*

$$\varphi^{\mathfrak{A}}_{\mathbf{a},\gamma} = \bigwedge\{\varphi^{\mathfrak{A}}_{\mathbf{a},\alpha}\colon \alpha < \gamma\}.$$

We collect two immediate properties of these formulae in the following proposition.

Proposition 2.2 *For every* $\mathbf{a}$*, we have*

(i) $\mathrm{qr}(\varphi^{\mathfrak{A}}_{\mathbf{a},\alpha}[\mathbf{x}]) = \alpha$

(ii) $\mathfrak{A} \models \varphi^{\mathfrak{A}}_{\mathbf{a},\alpha}[\mathbf{a}]$.

(iii) *If* $\beta < \alpha$*, we have* $\mathfrak{A} \models \forall \mathbf{x}\, [\varphi^{\mathfrak{A}}_{\mathbf{a},\alpha}[\mathbf{x}] \to \varphi^{\mathfrak{A}}_{\mathbf{a},\beta}[\mathbf{x}]]$

The sequence $\{\varphi^{\mathfrak{A}}_{\mathbf{a},\alpha} : \alpha < \omega_1\}$ provides increasing information about the structure $\mathfrak{A}$ as seen from the vantage point of the tuple $\mathbf{a}$. Put another way, they describe the isomorphism types of substructures of that are generated by extending a given tuple, one element at a time, up to the specified ordinal. In that sense, they are the direct generalization of the m-isomorphism types which performed this task for finite ordinals. Of course, the resulting formulae are no longer first order since they require infinitary conjunctions and disjunctions. Instead, they fall in $\mathscr{L}_{\infty\omega}$.

It remains to tie the Scott formulae with equivalence up to a certain quantifier rank in the logic $\mathscr{L}_{\infty\omega}$.

Theorem 2.3 *The following are equivalent.*

(i) $\mathfrak{A} \models \varphi^{\mathfrak{A}}_{\mathbf{b},\alpha}(\mathbf{a})$.

(ii) $(\mathfrak{A}, \mathbf{a}) \equiv_{\infty\omega,\alpha} (\mathfrak{B}, \mathbf{b})$.

We follow the sequence we have been for the entire chapter - first tying structural notions to equivalences in certain logics, and later to games. To accomplish this in the case of Scott formulae requires the notion of *rank* of a position in a game, which captures how "safe" player $\exists$ is in the game at that position, or put another way, how many more rounds she can play without being defeated.

Definition 2.4 *By a* position *of a game, we mean a tuple of possible moves in the game, starting at the beginning of the game. Let* $(\mathbf{a}, \mathbf{b})$ *be such a position. Then its* rank *is defined as follows.*

- $\mathrm{rank}(\mathbf{a}, \mathbf{b}) \geq 0$ *if none of the initial portions of the play* $(\mathbf{a}, \mathbf{b})$ *is a win for* SPOILER. *Namely, if* DUPLICATOR *has not lost by any previous round.*
- *For successor ordinals of the form* $\alpha + 1$, *the* $\mathrm{rank}(\mathbf{a}, \mathbf{b}) \geq \alpha + 1$ *if no matter how* SPOILER *moves,* DUPLICATOR *can respond in a way that the resulting position has rank* $\geq \alpha$.
- *For limit ordinals* δ, *the* $\mathrm{rank}(\mathbf{a}, \mathbf{b}) \geq \delta$ *if for all* $\alpha < \delta, \mathrm{rank}(\mathbf{a}, \mathbf{b}) \geq \alpha$.

Using this notion of rank of positions, we can now define an equivalence relation on structures of the same vocabulary.

Definition 2.5 *Let α be an ordinal. Two σ-structures are α-equivalent if the initial position has rank $\geq \alpha$. We denote this equivalence relation by $\sim^{\alpha}$.*

In other words, $\mathfrak{A} \sim^{\alpha} \mathfrak{B}$ if the DUPLICATOR can guarantee that she can play $G_{\omega}(\mathfrak{A}, \mathfrak{B})$ for at least α rounds without being defeated.

The following theorem is important in setting up the relation between the game $G_{\omega}(\mathfrak{A}, \mathfrak{B})$ and the language $\mathscr{L}_{\infty\omega}$.

Theorem 2.6 *For every ordinal α and $n < \omega$, there is a formula for every n-tuple $\mathbf{b}$ from B such that if $\mathbf{a}$ is a n-tuple from A, then the following are equivalent:*

(i) $\mathfrak{A} \models \varphi^{\mathfrak{A}}_{\mathbf{b},\alpha}(\mathbf{a})$.

(ii) *The position $(\mathbf{a}, \mathbf{b})$ has rank $\geq \alpha$ in $G_{\omega}(\mathfrak{A}, \mathfrak{B})$.*

We can now tie various equivalent ways of expressing the same condition together.

Theorem 2.7 *The following are equivalent:*

(i) $\mathfrak{A} \models \varphi^{\mathfrak{A}}_{\mathbf{b},\alpha}(\mathbf{a})$.

(ii) $(\mathfrak{A}, \mathbf{a}) \equiv_{\infty\omega,\alpha} (\mathfrak{B}, \mathbf{b})$.

(iii) $(\mathfrak{A}, \mathbf{a}) \sim^{\alpha} (\mathfrak{B}, \mathbf{b})$.

The sequence $\{\varphi^{\mathfrak{A}}_{\mathbf{a},\alpha} : \alpha < \omega_1\}$ cannot provide us new information beyond a certain ordinal, and that ordinal must be countable since the formulae $\varphi^{\mathfrak{A}}_{\mathbf{a},\alpha}[\mathbf{x}]$ define smaller and smaller subsets of A^n for larger and larger α. Thus, for each $\mathbf{a}$, there is a countable ordinal $\gamma_{\mathbf{a}}$ such that

$$\mathfrak{A} \models \forall \mathbf{x}\, [\varphi^{\mathfrak{A}}_{\mathbf{a},\gamma_{\mathbf{a}}}(\mathbf{x}) \rightarrow \varphi^{\mathfrak{A}}_{\mathbf{a},\beta}(\mathbf{x})] \text{ for all countable } \beta.$$

Definition 2.8 *The ordinal* $\gamma_{\mathbf{a}}$ *is called the* local Scott rank *of* $\mathbf{a}$*. The supremum of the local Scott ranks over all tuples in* $\mathfrak{A}$ *is an ordinal called the* Scott rank *of* $\mathfrak{A}$ *and is itself countable.*

Definition 2.9 *Let* γ *be the Scott rank of* $\mathfrak{A}$*. The* Scott sentence *of* $\mathfrak{A}$ *is the statement*

$$\Phi^{\mathfrak{A}} \triangleq \varphi^{\mathfrak{A}}_{\emptyset,\alpha} \wedge \bigwedge \{\forall \mathbf{x}\,[\varphi^{\mathfrak{A}}_{\mathbf{a},\gamma}(\mathbf{x}) \rightarrow \varphi^{\mathfrak{A}}_{\mathbf{a},\gamma+1}(\mathbf{x})] \colon \mathbf{a} \in A\}.$$

Note that the conjunction is taken over all tuples (of every length $< \omega$).

Proposition 2.10 *If* $|A| = \aleph_0$*, the Scott sentence of* $\mathfrak{A}$ *has only countably many conjunctions, and is therefore a sentence of* $\mathscr{L}_{\omega_1\omega}$*.*

The Scott sentence is constructed in manner as to encode the information required for a back-and-forth system to be constructed. More precisely, we have the following theorem.

Theorem 2.11 *Let* $\mathfrak{A}$ *and* $\mathfrak{B}$ *be two structures. The following are equivalent:*

(i) $\mathfrak{B} \models \Phi^{\mathfrak{A}}$,

(ii) *There is a back-and-forth system between* $\mathfrak{A}$ *and* $\mathfrak{B}$*, and*

(iii) $\mathfrak{A}$ *and* $\mathfrak{B}$ *are partially isomorphic.*

All of the previous discussion (except the Lemma 2.10) was agnostic to the cardinality of the structures. The remarkable fact about Scott sentences is that they characterize *countable* structures up to isomorphism.

Theorem 2.12 *Let* $\mathfrak{A}$ *and* $\mathfrak{B}$ *be two structures with* $|A| = |B| = \aleph_0$*. The following are equivalent:*

(i) $\mathfrak{B} \models \Phi^{\mathfrak{A}}$,

(ii) $\mathfrak{A} \cong \mathfrak{B}$.

Corollary 2.13 *Two countable structures are partially isomorphic if and only if they are isomorphic.*

Remark 2.14 *It is possible for a countable structure to be partially isomorphic to an uncountable structure.*

Remark 2.15 *The use of back-and-forth arguments to prove that two countable structures are isomorphic dates back to Cantor who used the argument to show that any countable dense lineary ordered set without endpoints is isomorphic to the rational numbers with their natural ordering.*

References

[Bar73] Jon Barwise. Back and forth through infinitary logic. In *Studies in model theory*, pages 5–34. MAA Studies in Math., Vol. 8. Math. Assoc. Amer., Buffalo, N.Y., 1973.

[Hod93] Wilfrid Hodges. *Model theory*, volume 42 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 1993.

[Mar10] David Marker. *Model theory: an introduction*, volume 217 of *Graduate texts in Mathematics*. Springer-Verlag, New York, 2010.

[Ros82] Joseph G. Rosenstein. *Linear orderings*, volume 98 of *Pure and Applied Mathematics*. Academic Press Inc. [Harcourt Brace Jovanovich Publishers], New York, 1982.