A TROJAN HORSE FOR LOGIC

JEREMI FORMAN-DURAÑONA

Abstract. Logic has a high barrier to entry for mathematics students. This paper aims to give a more concrete and natural introduction to some powerful ideas in logic by considering graphs and building intuition with the Ehrenfeucht game. A theorem about zero-one laws for graphs is used as a proving ground for the tools we develop.

Contents

1. Introduction

Logic is powerful. The notion of an inexpressibility or undecidability proof can send shivers down the spine: you can prove something is not provable?! Gödel's celebrated (dreaded?) incompleteness theorems are a prime example of logic having deep consequences for mathematics at large. However, the applications of logic to mathematics are not only in the negative – *you can't do this* – but also in the positive. We will give two sides to a beautiful example of this in random graph theory.

There is a problem with logic. Because logic is concerned with meta-mathematical structure, there are a lot of subtle concepts early on. What's more, the powerful theorems take a good amount of effort to prove and the generality of logic can make it unintuitive. All these factors can easily leave a student disillusioned with the subject.

If the reader finds this situation relatable, then hopefully this paper will prove helpful. If the reader has not experimented with logic at all, then hopefully this paper can help mitigate picture that has been painted.

Some familiarity with elementary probability would be helpful, but this paper should be more or less accessible to anyone comfortable with proofs.

2. Graphs and Zero-One Laws

Logic can be applied to study many structures. Our aim in this paper is to develop and demonstrate some basic but powerful tools from logic in the concrete environment of graphs to build intuition. So, before we start, it is essential we are on the same page about graphs.

Notations 2.1. If V is a set with $n := |V|$ and $k \leq n$ is an integer, then by $\binom{V}{k}$ we mean the set of all subsets of V of size k. We define "n choose k", written $\binom{n}{k}$ as $\left|\binom{V}{k}\right|$.

Definition 2.2. A *graph* $G = (V, E)$ is a set V called the *vertices* together with $E \subseteq \binom{V}{2}$ called the *edges*. If G is a graph, then by $V(G)$ we mean the vertex set of G and similarly for $E(G)$. The *order* of a graph is the cardinality of its vertex set. The *size* of a graph is the cardinality of its edge set. If $v, w \in V$ and $\{v, w\} \in E$ then we say v is *adjacent* or *incident* to w and write $v \sim w$. Unless otherwise stated, the vertex set of a graph is assumed to be finite. If v is a vertex and X is a set of vertices, by $v \sim X$ we mean $\forall x \in X, v \sim x$. Distance in a graph is defined to be the length of the shortest path, i.e. the number of edges between the endpoints.[1]

Remark 2.3. There can only be one edge between any two vertices, and no edges from a vertex to itself

We must make an important clarifying point. For graphs $G = (V, E)$ and $H = (W, F)$, we write $G = H$ and say G is the same as H if $V = W$ and $E = F$. However, often we are concerned with *isomorphic* graphs. That is, $G \cong H$ if there is a bijective map $\psi : V \to W$ such that $x \sim y \iff \psi(x) \sim \psi(y)$.

Another important notion in graph theory is that of a subgraph, and in particular, an induced subgraph. Here we define both.

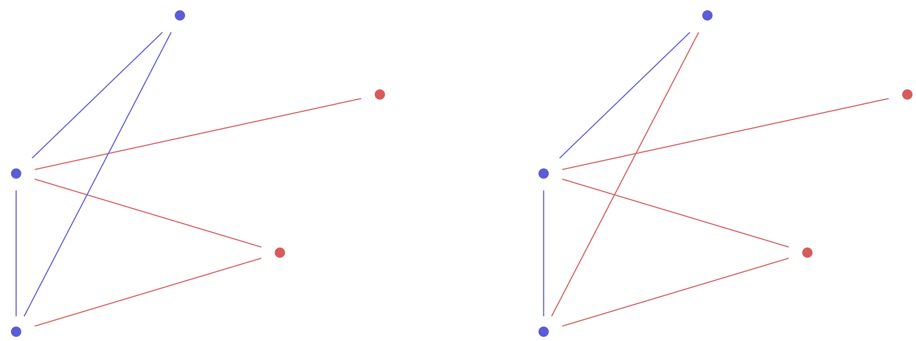
Definition 2.4. Let $G = (V, E)$. A graph $G' = (V', E')$ where $V' \subseteq V$ and $E' \subseteq E$ is called a *subgraph* of G. If this is the case and moreover, E' contains all edges in E involving only those vertices in V' then G is called an *induced subgraph*. We often denote the induced subgraph G' by $G[V']$. If $G = (V, E)$ and we want to consider the subgraph of G induced by vertices $V' = \{v_1, ..., v_k\}$ then we may merely write $G[v_1, ..., v_k]$ instead of $G[\{v_1, ..., v_k\}]$.

Another way to think about induced subgraphs is, given a subset V' of vertices from G, forget all vertices not in V' and the edges involving them. The graph we get in the end is an induced subgraph.

Example 2.5 (Induced Subgraph)**.** On the left and right is the same graph, G. The subgraph in blue on the *left* is induced, whereas the subgraph in blue on the

[1] There is a distinction between "paths" and "walks" etc. but we will not be concerned with this.

right is not.



There are a few of important graphs that receive enough attention to merit their own special notation.

- The *path* graph $P_n = (V, E)$ is defined by $V = \{x_0, ..., x_n\}$ and $E = \{\{x_0, x_1\}, \{x_1, x_2\}, ..., \{x_{n-1}, x_n\}\}$ where the x_i are distinct.
- The *cycle* graph $C_n = (V, E)$ is the same as P_n except also $x_n \sim x_0$.
- The *complete* graph $K_n = (V, E)$ is defined by $V = \{x_0, ..., x_n\}$ where the x_i are distinct and moreover for all $i \neq j$, $x_i \sim x_j$.

If x, y are vertices in a graph G, the *distance* between x and y, denoted $d(x, y)$, is the number of edges in the shortest path between x and y.

We'll also be interested in "properties" of graphs. Here are two examples.

- The *diameter* of a graph is the maximum distance between any two vertices in the graph. Having diameter equal to k is a property.
- The *chromatic number* of a graph G, denoted $\chi(G)$, is the least number of "colors" one can use to color the vertices of a graph so that no two adjacent vertices have the same color. Having $\chi(G) = k$ is a property.

We'll state a motivating result before proceeding to the next section. This theorem, which we prove in sections 4 and 6, is an example of what is called a "zero-one law." Informally, a given property P has a zero-one law if "for either almost all or almost no graphs, P holds."[2] This is not a vacuous statement: if the zero-one law holds, then it is *not* possible for the property P to hold for half of all graphs of every order. To be precise about the notion of "almost" always/never, we'll need a definition.

Definition 2.6. We denote by $G(n, p)$ the probability space on graphs on a set of n vertices where the probability of any two vertices being adjacent is set to p. For us, p will be a constant between zero and one. By $\mathcal{G} \sim G(n, p)$, we mean $\mathcal{G}$ is some elementary event in $G(n, p)$. One can think of $\mathcal{G}$ as a graph sampled at random from $G(n, p)$. Given a particular graph $G = (V, E)$ on those n vertices, the probability $\Pr[\mathcal{G} = G] = p^{|E|}(1-p)^{\binom{|V|}{2} - |E|}$.

Remark 2.7. To be clear, we are considering a set of n vertices and $G(n, p)$ is the set of all $2^{\binom{n}{2}}$ graphs together with a probability distribution on that set. In particular, there exist isomorphic graphs in $G(n, p)$ that are considered distinct elementary events.

[2] Zero-one laws show up not just for graph properties, but we will only consider a particular zero-one law for graphs.

Example 2.8. For simplicity, let's think of $G(4, 1/2)$ as the probability distribution over all graphs on $[4] \coloneqq \{k \in \mathbb{N} : k \leq 4\}$ as defined in Definition 2.6. If $\mathcal{G} \sim G(4, 1/2)$ then what is the probability that $\mathcal{G}$ is *equal* to the cycle graph C_4 where each vertex is "in order," i.e. $1 \sim 2 \wedge 2 \sim 3 \wedge 3 \sim 4 \wedge 4 \sim 1$? Since the probability of an edge existing is the same as a nonedge existing ($p = 1/2$) we simply get $(1/2)^{\binom{4}{2}} = (1/2)^{4(4-1)/2} = 1/2^6$. However, this is not the only cycle graph of order 4 that $\mathcal{G}$ could be. What then is the probability that $\mathcal{G} \cong C_4$? There are only 3 distinct orderings, not double counting symmetries. (For instance, the ordering $1, 2, 3, 4$ is considered the same as $4, 1, 2, 3$ or $4, 3, 2, 1$.) It follows that $\Pr[\mathcal{G} \cong C_4] = 3 \Pr[\mathcal{G} = C_4] = 3(1/2^6)$.

Now we are equipped to state the formal definition of a zero-one law for graphs.

Definition 2.9. Let ϕ be a property of graphs. Let $\mathcal{G}_n \sim G(n, p)$. We call the following a zero-one law:

$$\lim_{n\to\infty} \Pr[\mathcal{G}_n \models \phi] = 0 \text{ or } 1. \tag{2.10}$$

The symbol "$\models$", read "satisfies," simply means "has property." So $\Pr[\mathcal{G}_n \models \phi]$ is the probability that $\mathcal{G}_n$ has the property ϕ.

Typically we are interested in zero-one laws for classes of properties, not just one property: a single property satisfying (2.10) may as well be said to satisfy a "zero law" (or a "one law").

The reader should note that the limit in (2.10) may not even converge for a given property: consider the property "order of graph is odd." However, for many properties not only does the limit converge, but it also converges to either zero or one. If either of these is the case, we say such a property holds *almost never* or, respectively, *almost always*.

Example 2.11.

(1) Almost all graphs contain a triangle;
(2) almost no graphs are triangle-free;
(3) almost all graphs have diameter 2;
(4) almost all graphs are connected;
(5) for any $\epsilon > 0$, almost no graph $G = (V, E)$ with $n = |V|$ has chromatic number[3]
$$\chi(G) \leq \frac{\log(1/(1-p))}{2+\epsilon} \cdot \frac{n}{\log n}.$$

Exercise 2.12. Prove (1) in Example 2.11 in the sense that for $0 < p < 1$ a constant, and $\mathcal{G}_n \sim G(n, p)$,

$$\lim_{n\to\infty} \Pr[\mathcal{G}_n \text{ contains a triangle}] = 1.$$

Equivalently, prove (2).

The first three properties belong to a special class of properties called "first order." What makes a property first order or not is the subject of the next section. Essentially, they are properties that can be expressed in logic while obeying certain rules. We'll use the tools we develop to prove the following fact involving first-order properties.

[3]See [3].

Theorem 2.13. *Let ϕ be a first-order property and let $\mathcal{G}_n \sim G(n,p)$. Then we have the zero-one law*

$$\lim_{n\to\infty} \Pr[\mathcal{G}_n \models \phi] = 0 \text{ or } 1.$$

This theorem is quite surprising. As one may glimpse in Example 2.11 and will continue to see, the first-order language is quite expressive. Yet, the theorem holds!

The rules for deciding what properties are first order are simple, but they do not provide much understanding. In fact, the precise rules are nearly useless once one wants to prove something about first-order properties that isn't obvious. We'll get to the rules, but we'll start instead with a game that hopefully provides some intuition, utility in proofs, and most of all – fun.

3. Do you want to play a game?

Let's play a fun and seemingly innocent game. It's played on graphs.

Definition 3.1 (Ehrenfeucht Game)**.** The game is played by two players: Spoiler (he/him/his) and Duplicator (she/her/hers).[4] The game is played in k rounds and on two given graphs G, H. We denote the game fitting these parameters by $\texttt{EHR}(G, H, k)$. The game is played as follows. On the i-th round, Spoiler goes first and selects a vertex of one of the graphs. After Spoiler picks a vertex of one graph, Duplicator picks a vertex of the other graph. These vertices are labeled i_G and i_H respectively. If the induced graphs $G[1_G, ..., k_G], H[1_H, ..., k_H]$ are (order)-isomorphic then Duplicator wins. That is, Duplicator wins if for all i, j, we have that $i \sim j$ or $i = j$ in one graph iff it does in the other. Otherwise, Spoiler wins.

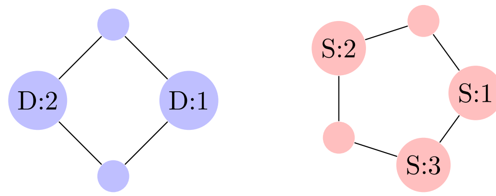
Remark 3.2. Note that a vertex that has already been selected can be selected again according to this definition. The game could equivalently be formulated forcing a new vertex to be chosen every round, however, our definition avoids mentioning some extra details in the ensuing proofs.

For an interactive sense of this game, Thomas Kern made a fantastic website where one can play against an AI or a friend.

Clearly only one player will win after k rounds, but a result from game theory states that given the parameters of G, H, and k, there exists a winning strategy for one of the players. In other words, one of the players has a forced win from the beginning. The proof may be found in [1]. While this winning strategy may be hard to find, its existence (for a particular player) is what we are interested in. Going forward, by Duplicator (resp. Spoiler) *wins*, we mean Duplicator (resp. Spoiler) has a *winning strategy*.

Example 3.3. Consider $\texttt{EHR}(C_4, C_5, k)$. What is the smallest $k > 0$ such that Spoiler can win? Clearly $k = 1$ is safe for Duplicator. One can exhaustively consider every scenario in $k = 2$ and again Duplicator has a winning strategy. Here is a winning strategy for Spoiler at $k = 3$:

[4]We follow the convention found in the literature of ascribing pronouns to the players – a measure of convenience when explaining strategies in the game.



S:i means Spoiler marked a given vertex i and similarly for Duplicator.

As a hint of what is to come, this game turns out to give an equivalent characterization of so-called first-order logic. For now, let's just play the game. Here is a rough observation.

Example 3.4. If G and H are isomorphic graphs, then no matter what k is, Duplicator wins $\texttt{EHR}(G, H, k)$. This is because the isomorphism provides her (Duplicator) with a safe move no matter what. If G and H are not isomorphic, then for any $k > \min\{|G|, |H|\}$, Spoiler wins $\texttt{EHR}(G, H, k)$. If $|G| = |H|$, then because they aren't isomorphic, Spoiler will uncover a difference. Suppose without loss of generality that $|G| > |H|$. Then Spoiler can just pick only new vertices on G and at some point Duplicator will have to pick a vertex of H she has already picked, which means there are $i \neq j$ such that $i_H = j_H$ but $i_G \neq j_G$: Spoiler wins.

What if there is a particular property on which G and H disagree? Can Spoiler always "prove" the difference? Consider the property of being 2-*colorable*, i.e. having chromatic number of 2 as defined on page 3.

Exercise 3.5. A graph is 2-colorable iff it does not contain an odd-cycle.

Proposition 3.6. *Duplicator wins* $\texttt{EHR}[C_n, C_{n+1}, k]$ *when* n *is sufficiently large.*

Proof. With s rounds remaining, we call two vertices *close* if they are a distance at most 2^{s-1} away from each other. Otherwise they are *far*. We call two positions equivalent if for all $i, j \leq k - s$ with s moves remaining, we have $d(i_G, j_G) = d(i_H, j_H)$ for close i, j. If a move is close to a previous move, it is called *inside*. Otherwise it is called *outside*. We then have to consider two cases to proceed by induction.

(Case 1: Inside) Let the positions be equivalent in C_n and C_{n+1} with s moves remaining. Suppose Spoiler moves inside with $x \in G$ (works similarly for H). Then for some i, $d(x, i_G) \leq 2^{s-1}$. Locally, C_n and C_{n+1} are isomorphic, so in the same orientation, Duplicator picks y so that $d(y, i_H) = d(x, i_G)$. We can check if the new positions are equivalent. Suppose y is close to y_j after the $(k-s)$th move. But then by the triangle inequality,

$$d(i_H, j_H) \leq d(i_H, y) + d(y, j_H) \leq 2^{s-1} + 2^{s-1} = 2^s$$

meaning i_H, j_H were already close with s moves remaining so by our inductive hypothesis $d(i_G, j_G) = d(i_H, j_H)$ with the correct orientation. It follows that the distance between x and its close vertices is the same between y and its close vertices, i.e. the positions remain equivalent.

(Case 2: Outside) Suppose the positions are equivalent and on the $(k-s)$th move, Spoiler picks an outside vertex. We'll assume Duplicator has to respond on C_n, but the argument is the same for C_{n+1}. We need to only demonstrate that for any s there exists an outside vertex for Duplicator (when $s = k$ every vertex is outside). In the spirit of what we'll do later, we'll do a probabilistic proof. We'll show there exists an n so that given any $1 \leq s \leq k-1$ rounds remaining, the probability of Duplicator picking an outside vertex is nonzero, so in particular there exists an outside vertex. We can equivalently show that the probability of Duplicator picking an inside vertex is less than one. Let y represent the vertex Duplicator picks in response. By the union bound, we have

$$\begin{aligned} \Pr[d(y, 1_H) \leq 2^{s-1} \vee \cdots \vee d(y, (k-s)_H) \leq 2^{s-1}] &\leq (k-s)\frac{[2 \cdot 2^{s-1} - 1]}{n} \\ &= (k-s)\frac{[2^s - 1]}{n}. \end{aligned} \tag{3.7}$$

If for every $1 \leq s \leq k-1$, we have $n > (k-s)[2^s - 1]$ then there always exists an outside vertex, and Duplicator's strategy succeeds. It turns out $s = k-1$ maximizes the expression, so n need only be greater than $2^{k-1} - 1$ for her to succeed. Note that though we imagine Duplicator picks a vertex at random for our argument, Duplicator does *not* play randomly: if there exists an outside vertex with which to respond, she will pick it. The random argument was a simple example of the "probabilistic method" – a staple of graph theory that is particularly handy for nonconstructive existence proofs. □

Remark 3.8. It is absolutely key that the notion of "closeness" converges to "adjacency" as $s \to 0$. If at each round, vertices i_G, j_G are close iff i_H, j_H are close then Duplicator wins.

Exercise 3.9. This is a slightly more involved exercise and is not necessary. If $n \leq 2^k + 1$ and $n < m$ then Spoiler wins $\texttt{EHR}(P_n, P_m, k+2)$. Moreover, if $n, m \geq 2^{k+1} + 1$ then Duplicator wins $\texttt{EHR}(P_n, P_m, k)$. *Hints:* For the Spoiler strategy: start with the endpoints of P_n. For the Duplicator strategy, refactor the proof of Proposition 3.6 and allow Spoiler the endpoints to start with.

It seems by now that if Duplicator has a winning strategy for a given $\texttt{EHR}(G, H, k)$ then G and H have *something* in common and that this is increasingly the case for higher k. To pinpoint this exactly, we need a few definitions and a powerful theorem.

Notice that given a fixed k, the wins for Duplicator form an equivalence relation on all graphs:

Definition 3.10. If Duplicator wins $\texttt{EHR}(G, H, k)$ then we write $G \equiv_k H$.

Notice $\equiv_k$ is reflexive: if two graphs are the same, Duplicator always has a safe move. Moreover, the order of the graphs in the game has no impact on the game so $\equiv_k$ is symmetric. We leave transitivity to the reader.

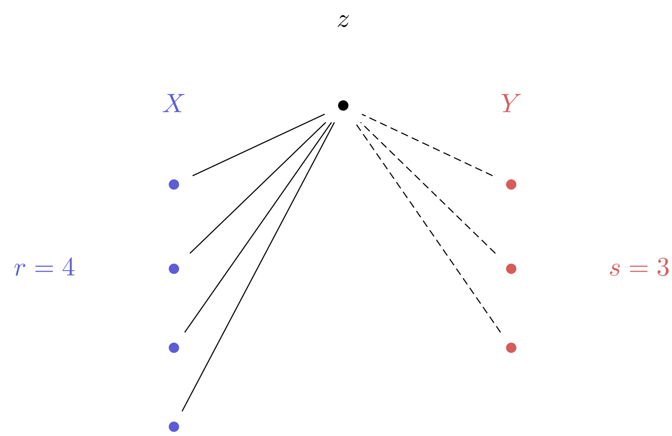
Exercise 3.11. Prove that if $F \equiv_k G$ and $G \equiv_k H$ then $F \equiv_k H$.

Since $\equiv_k$ is an equivalence relation, it is only natural to consider its equivalence classes. In the literature, the equivalence classes are called *k-Ehrenfeucht values.* This is needlessly confusing. We will simply either stomach the extra syllables and call the equivalence classes of $\equiv_k$, *k-Ehrenfeucht equivalence classes* or $\equiv_k$ classes.

What does a given $\equiv_k$ class look like? This is a tricky question to answer.[5] However, here is an instructive (and luckily tractable) example: 3.15. We'll need a definition, which we will reuse later.

[5]The Trakhtenbrot-Vaught Theorem places logical bounds on a complete description ([7]).

Definition 3.12. A graph G is said to have the $A_{r,s}$ *extension property* if for any disjoint subsets X and Y of the vertex set $V = V(G)$ where $|X| = r$ and $|Y| = s$ there exists a *witness* $z \in V \setminus (X \cup Y)$ such that $z \sim X$ and $z \nsim Y$.



Definition 3.13. A graph G is said to have the k-*Alice's Restaurant property* if it has more than k vertices and if for all r, s such that $r + s \leq k$, G has the $A_{r,s}$ extension property.

Remark 3.14. If for all k, G has the k-Alice's Restaurant property then it is simply said to have the *Alice's Restaurant property.* Note that such a graph must be infinite. Peter Winkler coined the term after Arlo Guthrie's whimsical and beloved (roughly 18-minute-long) song of the same name where it is stated that "You can get anything you want at Alice's Restaurant." Indeed as we will find in Section 6, the property is aptly named.

Example 3.15. The set of all graphs possessing the $(k-1)$-Alice's Restaurant property form an $\equiv_k$ class.

Proof. Let G and H possess the $(k-1)$-Alice's Restaurant property. The property guarantees that while $s \leq k-1$, Duplicator has a safe $(s+1)$st move. This means Duplicator has a safe move up to round k. □

We're well on our way to drawing a connection between Ehrenfeucht games and first-order logic. So far, we have really only been considering "semantics," i.e. graphs *having* particular properties. To continue, we must also be concerned with how we *express* those properties with language: syntax. The first-order properties we are concerned with are those that can be expressed according to the rules in the following definition.

Definition 3.16. A formula (on graphs) is called *first order* if it can be expressed in a valid combination of

- finitely many variable symbols, which for us stand for vertices of a graph (e.g. v, w, x, y, etc.);
- (binary) relation symbols $=$ and $\sim$;
- the usual logical connectives $\wedge, \vee, \neg$ ("and," "or," and "not");

- and the quantifiers $\forall$ and $\exists$, which are allowed to quantify over all vertices of the graph.[6]

If a property can be expressed by a first-order formula, then we say the property is first order.

The final bullet point is crucial. For a formula to be first order, it cannot quantify over sets of vertices. For instance, "for all paths P in G, ..." is not allowed. If one could express arbitrarily quantifying over paths in a graph by the rules in Definition 3.16 then that would be alright. However, it turns out that is not possible.

Showing that a property is first order is typically not difficult: write it down following the rules. However, showing that a property is not first order seems more difficult. It's as though we took a subset of the English vocabulary as well as a subset of its grammatical rules and ask what statements are *not* possible to form. Difficult to say.

Exercise 3.17. Write down a first-order formula encoding the property of containing a triangle.

Example 3.18. Not only is containing a triangle a first-order property, but the property ϕ of containing any given graph H as an induced subgraph is also first order. We'll use some shorthand, but it stands in only for valid applications of the rules in Definition 3.16. For notational convenience, we'll take an ordering of the vertices of H: $x_1, ..., x_n$. However, note that the ordering is arbitrary. Let $N(x_i)$ denote the set of vertices adjacent to x_i and $N(x_i)^c$ the set of non-adjacent vertices.

$$(\exists x_1)\cdots(\exists x_n)\left[\left(\bigwedge_{i\neq j} x_i \neq x_j\right) \wedge \left(\bigwedge_i x_i \sim N(x_i)\right) \wedge \left(\bigwedge_i x_i \not\sim N(x_i)^c\right)\right]$$

The formula says that there are n distinct vertices and that they have precisely those adjacencies in H.

Definition 3.19. If a property ϕ is true of a graph G, we write $G \models \phi$ and say G *models* or *satisfies* ϕ. We use the same notation if we interpret a first-order sentence as its corresponding property, i.e. if the property expressed by the sentence is true of the given graph.

Recall the earlier observation immediately below Exercise 3.9 that a win for Duplicator corresponds to a similarity between the graphs on which the game was played, and moreover that higher k indicates more similarity. To make this more clear, we'll need the notion of quantifier depth.

Definition 3.20. Informally, the *quantifier depth* of a formula is the largest number of nested quantifiers in the formula. Formally, we inductively define the quantifier depth (QD) of a formula ϕ as follows:

(1) if ϕ has no quantifiers, then the QD of ϕ is 0;
(2) if $\phi = (\forall x)\psi(x)$ or $\phi = (\exists x)\psi(x)$ then the QD of ϕ is one more than the QD of ψ;
(3) if $\phi = \psi \wedge \pi$ or if $\phi = \psi \vee \pi$ then the QD of ϕ is the maximum of the QD's of ψ and π;

[6] We can formally define what "valid" means, but a mathematics student should have a good idea from experience.

(4) if $\phi = \neg\psi$ then the QD of ϕ is that of ψ.

Example 3.21. The formula $(\forall x)(\forall y)(\exists z)(x \sim y \implies x \sim z \wedge y \sim z)$ has QD of 3.

Finally, our connection between the Ehrenfeucht game and first-order logic is made precise by the following big theorem. Recall that in Example 3.15, it is a particular property that characterizes a given $\equiv_k$ class. It turns out that the Alice's Restaurant property is first order. Moreover, it turns out that this is true more broadly: for any $\equiv_k$ class α, there exists a first-order property that characterizes α. This is the first part of the following theorem.

Theorem 3.22.

(1) To each $\equiv_k$ class α, there exists a first-order property ϕ of quantifier depth k such that $G \in \alpha \iff G \models \phi$.
(2) $G \equiv_k H$ if and only if G and H satisfy the same set of first-order sentences of quantifier depth k.

Note that (2) is equivalent to G and H satisfying the same set of first-order sentences of quantifier depth *up to and including* k.

We will need a definition and a lemma to prove this theorem. In return, we prove a slightly stronger result: Theorem 3.30.

Definition 3.23. We will denote by $\mathtt{EHRC}[k]$ the set of all of the $\equiv_k$ classes. The following framing is useful in proofs. If s rounds have already been played in $\mathtt{EHR}(G, H, k)$ and Duplicator has a winning strategy given the current board and the remaining $k - s$ rounds then we write

$$(1_G, ..., s_G) \equiv_k (1_H, ..., s_H). \tag{3.24}$$

By $\mathtt{EHRC}[k, s]$, we denote the set of equivalence classes induced by $\equiv_k$ on graphs together with s marked vertices. Let it be clear that $\mathtt{EHRC}[k, s]$ is *not* a partition of the set of graphs, but rather is a partition of the set of $(s+1)$-tuples, $(G, 1_G, ..., s_G)$, of a graph G and s labels of its vertices. We may call an equivalence class $\alpha \in \mathtt{EHRC}[k, s]$ an $\equiv_k^s$ *class.*

Exercise 3.25. Check $\equiv_k^s$ for $1 \le s \le k$ is also an equivalence relation.

To prove Theorem 3.22 we will need to know that $\mathtt{EHRC}[k]$ is finite for any given k.

Lemma 3.26. *For all k and s, $|\mathtt{EHRC}[k]|$ and $|\mathtt{EHRC}[k, s]|$ are finite.*

Proof. We claim that

(1) $|\mathtt{EHRC}[k, k]|$ is finite and that
(2) $|\mathtt{EHRC}[k, s]| \le |\mathtt{EHRC}[k, s + 1]|$.

The lemma follows from these two claims.

(1) By construction, $(1_G, ..., k_G) \equiv_k (1_H, ..., k_H)$ if the two induced graphs $G[1_G, ..., k_G]$ and $H[1_H, ..., k_H]$ are isomorphic. It follows that $|\mathtt{EHRC}[k, k]|$ is equal to the cardinality of the set of isomorphism classes of graphs on k vertices. A rough upper bound to this number is just the number of possible graphs on k vertices: $2^{\binom{k}{2}}$.

(2) We'll just use the rough bound from (1). We will make a surjective map $\texttt{EHRC}[k, s+1] \to \texttt{EHRC}[k, s]$. This proves (2). When dealing with equivalence classes is is helpful to instead consider representative members of each of those equivalence classes. The following $\texttt{EHRC}[k, s+1] \to \texttt{EHRC}[k, s]$ map is surjective:

$$(G, 1_G, ..., s_G, (s+1)_G) \mapsto (G, 1_G, ..., s_G)$$

where $(G, 1_G, ..., s_G, (s+1)_G)$ is a representative member of a $\equiv_k^{s+1}$ class. Thus to each $\alpha \in \texttt{EHRC}[k, s]$, there is associated at least one $\beta \in \texttt{EHRC}[k, s+1]$. It follows that $|\texttt{EHRC}[k, s]| \leq |\texttt{EHRC}[k, s+1]|$ and we have proved (2).

□

Though we will only use Lemma 3.26, it is possible to prove the following "tower-style" inequality.

Theorem 3.27. *Define* $T : \mathbb{N} \cup \{0\} \to \mathbb{N}$ *by* $T(0) = 1$ *and* $T(k+1) := 2^{T(k)}$. *Define* $\log^* : \mathbb{N} \to \mathbb{N} \cup \{0\}$ *by* $\log^*(n)$ *is the smallest* k *such that* $n \leq T(k)$.

For $k \geq 20$,

$$T(k-2) \leq |\texttt{EHRC}[k]| \leq T(k+2+\log^* k)$$

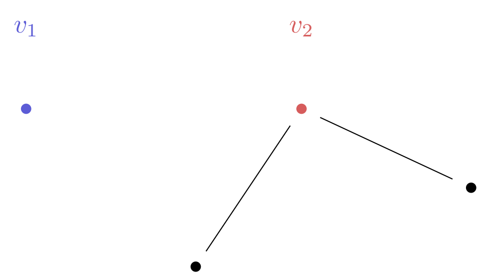
Proof. Proof found under Theorem 2.2.2. of [7]. □

This tower theorem goes to show that the $\equiv_k$ classes, though finite, are rather large.

We'll need to restate Theorem 3.22 in the slightly stronger form that we'll prove, but to understand it we need a definition.

Definition 3.28. In a logical expression, variables can appear with a quantifier: $(\forall x)(....)$ or $(\exists x)(....)$. Such variables are called *bounded* variables. We call variables that are not bounded *free* variables. Notice that the truth value of an expression involving free variables cannot be evaluated. However, if we refer to particular objects (vertices) by "plugging in" values into the variable "spots" the expression can be evaluated.

Example 3.29. The property $\phi(v) := (\exists x)(v \sim x)$ can only be assigned a truth value once v is "plugged in." For instance in the figure below, $\phi(v_1)$ is false, whereas $\phi(v_2)$ is true.



Theorem 3.30 (Stronger Form of 3.22)**.** *For all* $k \geq 1$ *and all* $0 \leq s \leq k$,

(1) To each $\equiv_k^s$ *class (i.e. each* $\alpha \in \texttt{EHRC}[k, s]$*) there is a first-order formula of quantifier depth* $k - s$ *and* s *free variables* $\phi(x_1, ..., x_s)$ *such that the* $(G, 1_G, ..., s_G)$*'s pertaining to that particular* $\equiv_k^s$ *class are precisely those for which* $\phi(1_G, ..., s_G)$ *is true.*

(2) $(G, 1_G, ..., s_G) \equiv_k (H, 1_H, ..., s_H)$ *if and only if* G, H *together with their selected vertices have the same truth value for all first-order formulae of quantifier depth* $k - s$ *with* s *free variables plugging in* $1_G, ..., s_G$ *(resp.* $1_H, ..., s_H$*).*

This is a big theorem. Let's unpack it. Item (1) states (as in 3.22) that the $\equiv_k$ classes are completely determined by a single property. We've seen an example of this in Example 3.15. The ideas of this proof are simpler than the notation suggests. We'll try to include some intuition amid the technical details.

Proof of Theorem 3.30.
We prove the theorem following a strategy of reverse induction. We will prove the following cases:

(a) (1) in the case that $s = k$;
(b) (2) in the case that $s = k$;
(c) (1) for some non-negative $s < k$ under the inductive hypothesis that the result holds for $s + 1$;
(d) (2) for s assuming (1) for s and also the inductive hypothesis of the result holding for $s + 1$.

Proof of (a). When $s = k$, $(G, 1_G, ..., k_G) \equiv_k (H, 1_H, ..., k_H)$ iff the induced subgraphs $G[1_G, ..., k_G]$ and $H[1_H, ..., k_H]$ are not only isomorphic, but the same. That is, the induced subgraphs are not only isomorphic but $1_G, ..., k_G$ gives the same ordering to the vertices as $1_H, ..., k_H$. It follows that for $s = k$, the formula characterizing the $\equiv_k^s$ classes would just list the adjacencies among the $1_G, ..., k_G$. This would be a formula of $s = k$ variables and $k - s = 0$ quantifiers. □

Proof of (b). ($\implies$) Suppose $(1_G, ..., k_G) \equiv_k (1_H, ..., k_H)$ and there exists a first-order $\phi(x_1, ..., x_k)$ of QD $= k - s = 0$ such that G and H with their selected vertices have different truth values for ϕ. Since the QD of ϕ is 0, ϕ is just some boolean combination of the atomic "$x_i \sim x_j$" and "$x_i = x_j$" expressions. Thus G and H with their selected vertices disagree on some "$x_i \sim x_j$" or "$x_i = x_j$" expression. But this is a contradiction, since $G[1_G, ..., k_G]$ and $H[1_H, ..., k_H]$ are order-isomorphic. It follows that $(G, 1_G, ..., k_G)$ and $(H, 1_H, ..., k_H)$ must have the same truth value for all first-order ϕ of QD $= 0$ and k free variables.

($\impliedby$) If $(G, 1_G, ..., k_G)$ and $(H, 1_H, ..., k_H)$ agree on all $\phi(x_1, ..., x_k)$ of QD $= 0$ then in particular they agree on the formula listing the adjacencies or equalities. In other words, $(G, 1_G, ..., k_G)$ and $(H, 1_H, ..., k_H)$ are order-isomorphic so in particular, $(1_G, ..., k_G) \equiv_k (1_H, ..., k_H)$. □

For (c) we must construct a characteristic formula for each $\alpha \in \texttt{EHRC}[k, s]$. The idea is to leverage our inductive hypothesis. We can "push up" one level to $\texttt{EHRC}[k, s + 1]$ and construct a formula using the characteristic formulae of the $\beta \in \texttt{EHRC}[k, s + 1]$. The way we do this is similar to Lemma 3.26: we select another vertex x in some $(G, 1_G, ..., s_G) \in \alpha$ and see which characteristic A_β are satisfied by $(G, 1_G, ..., s_G, x)$.

Proof of (c). By our inductive hypothesis, each $\beta \in \texttt{EHRC}[k, s + 1]$ has a characteristic $A_\beta(x_1, ..., x_{s+1})$ of QD $k - s - 1$. We wish to produce a similar characteristic property for any $\alpha \in \texttt{EHRC}[k,s]$. Let $\texttt{YES}[\alpha]$ be the set of β such that for a representative member $(G, 1_G, ..., s_G)$ of α, the formula

$$(\exists x) A_\beta(1_G, ..., s_G, x) \tag{3.31}$$

holds. Similarly, define $\texttt{NO}[\alpha]$ as the set of β for which (3.31) does not hold. We claim the following formula, denoted A_α, characterizes α:

$$\left[\bigwedge_{\beta \in \texttt{YES}[\alpha]} (\exists x)(A_\beta(x_1, ..., x_s, x))\right] \wedge \left[\bigwedge_{\beta \in \texttt{NO}[\alpha]} (\neg \exists x)(A_\beta(x_1, ..., x_s, x))\right]. \tag{3.32}$$

Suppose $(G, 1_G, ..., s_G) \in \alpha$ then it satisfies A_α by construction. We must now show that $(G, 1_G, ..., s_G)$ satisfying A_α implies $(G, 1_G, ..., s_G) \in \alpha$. Suppose $(G, 1_G, ..., s_G)$ and $(H, 1_H, ..., s_H)$ both satisfy A_α and that $(H, 1_H, ..., s_H)$ is in α. Suppose Spoiler moves in G, though the argument works if he moves in H. Whatever $(s+1)$st move Spoiler takes, by construction he will wind up in some $\beta \in \texttt{YES}[\alpha]$. Since there exists an $x \in H$ so that H satisfies A_β, by induction $(H, 1_H, ..., s_H, x) \in \beta$ as well. In other words, Duplicator wins. Since Duplicator wins no matter what $(s+1)$st move Spoiler makes, we have $(1_G, ..., s_G) \equiv_k (1_H, ..., s_H)$. It follows that since H with its selected vertices is in α that $(G, 1_G, ..., s_G)$ is also in α and we are done. □

Proof of (d). ($\Longleftarrow$) Suppose for all first-order ϕ of QD $k-s$ and s free variables, ϕ is satisfied by $(G, 1_G, ..., s_G)$ iff it is satisfied by $(H, 1_H, ..., s_H)$. Then in particular they both satisfy the same A_α and thus $(1_G, ..., s_G) \equiv_k (1_H, ..., s_H)$.

($\Longrightarrow$) Suppose $(G, 1_G, ..., s_G)$ and $(H, 1_H, ..., s_H)$ are $\equiv_k$ equivalent: Duplicator wins. Furthermore assume they disagree on a formula ϕ of QD $k-s$ and s free variables. We rewrite the formula as $\phi = (\exists x)\psi(x_1, ..., x_s, x)$. Such an x exists in G, but not in H, so by our inductive hypothesis $(1_G, ..., s_G, x) \not\equiv_k (1_H, ..., s_H, y)$ for any $y \in H$: Spoiler wins. We thus have a contradiction and $\equiv_k^s$ equivalence implies agreement on all first-order ϕ of QD $k-s$ and s free variables. □

Now that we have the big Theorem 3.30 proved we can take advantage of the power of the Ehrenfeucht game. Earlier it was noted that proving when a given property is *not* first order was not super clear. The following corollary of our big theorem gives a methodology for proving such results:

Corollary 3.33. *Let ϕ be a property. If for all k, there exist graphs G and H such that $G \equiv_k H$, but $G \models \phi$ and $H \not\models \phi$, then ϕ must not be first order.*

Proof. By part (2) of Theorem 3.30, if for all k, $G \equiv_k H$, then G and H agree on all first-order formulas. If for some ϕ, $G \models \phi$ but $H \not\models \phi$, then it follows ϕ must not be first order. □

What does a non-expressability proof using Ehrenfeucht look like? We've already seen one! Considered in light of Corollary 3.33, Proposition 3.6 is a proof that the property of being 2-colorable is not first order. The result can be generalized to k-colorability, and if the reader is interested, they are urged to continue in chapter 3 of [7]. The strategy we used in Proposition 3.6 can be applied to more properties.

A fun note to reiterate: recall from Example 2.11 that 2-connectedness (i.e. for any two vertices, there exists a third to which they are both adjacent) *is* first order. Moreover, almost all graphs are 2-connected, so in particular almost all graphs are connected. However, connectedness itself is not first order as can be proved in Exercise 3.34.

Exercise 3.34. Let G be a cycle on n vertices. Let H be two disjoint cycles on n vertices each respectively. Apply the same strategy from Proposition 3.6 to show that connectedness is not a first-order property.

The applications of the Ehrenfeucht game don't end with expressability results. In the next section, we turn to random graphs and the zero-one law, Theorem 2.13, we mentioned early in the paper for a surprising application.

4. Ehrenfeucht Applied to the Zero-One Law

A few pages have gone by since $G(n, p)$ was defined. Feel free to recall the definition: 2.6. We'll also restate Theorem 2.13 here so it's handy.

Theorem 4.1 (Zero-One Law for First-Order Properties)**.** *Let ϕ be a first-order property and let $\mathcal{G}_n \sim G(n, p)$ with a constant $0 < p < 1$. Then*

$$\lim_{n\to\infty} \Pr[\mathcal{G}_n \models \phi] = 0 \text{ or } 1. \tag{4.2}$$

We'll build up to the theorem by first seeing what we can learn by considering a typical Ehrenfeucht game on random graphs. Note that the players do *not* play randomly. The players play perfectly, it is only the graphs on which they play that are random. It may seem counterintuitive for Duplicator to typically win – *shouldn't two random graphs be pretty different?* – but, that is exactly what happens. In fact, that should seem reasonable after some thought: for a fixed k, the larger the graphs, the likelier it should be that Duplicator finds isomorphic induced subgraphs.

We provide a strategy for Duplicator that works for almost all pairs of graphs. Recall from Example 3.15 that the $(k-1)$-Alice's Restaurant property characterizes an $\equiv_k$ class. The strategy rests on the following fact.

Proposition 4.3. *The $(k-1)$-Alice's Restaurant property holds almost surely.*

Proof. It suffices to show that for arbitrary r, s, the $A_{r,s}$ extension property holds almost surely. Given fixed subset X of cardinality r and a disjoint fixed subset Y of cardinality s, let `NoZ` be the property that there does *not* exist a witness given the sets X and Y.

$$\Pr[\texttt{NoZ}] = (1 - p^r(1-p)^s)^{n-(r+s)}$$

Ranging over all fixed X and S, we have by the union bound

$$\Pr[\neg A_{r,s}] \leq \binom{n}{r}\binom{n-r}{s}(1 - p^r(1-p)^s)^{n-(r+s)} \to 0 \text{ as } n \to \infty.$$

Notice the binomial coefficients give a polynomial in n, whereas the rest of the bound decays exponentially with n. □

Now the almost sure strategy for Duplicator follows quite cleanly.

Lemma 4.4 (Almost Sure Strategy)**.** *Duplicator has a strategy that almost always works, i.e. if $\mathcal{G}_n \sim G(n, p_n)$ and $\mathcal{G}_m \sim G(m, p_m)$ then*

$$\lim_{n,m\to\infty} \Pr[\mathcal{G}_n \equiv_k \mathcal{G}_m] = 1.$$

Proof of Almost Sure Strategy. Fix k. In Example 3.15, we saw that graphs satisfying $(k-1)$-Alice's Restaurant property belong to the same $\equiv_k$ class. Since almost every pair of graphs $\mathcal{G}_n, \mathcal{G}_m$ belong to that $\equiv_k$ class, Duplicator must win on almost every pair of graphs. □

It should seem natural that Duplicator winning on almost every pair of graphs should have some logical consequences for the zero-one law. All we need to show is that it follows there is a zero-one law, but we'll go one further and prove that a zero-one law also implies Duplicator has an almost sure strategy. Because of this equivalent formulation of the zero-one law, we call the following a "bridge theorem."

Theorem 4.5 (Bridge Theorem)**.** *Let $\mathcal{G}_n \sim G(n,p)$ and $\mathcal{G}_m \sim G(m,p')$. There is a zero-one law for first-order properties iff for every $k \in \mathbb{N}$,*

$$\lim_{m,n\to\infty} \Pr[\mathcal{G}_n \equiv_k \mathcal{G}_m] = 1. \tag{4.6}$$

Proof of Bridge Theorem. First we prove that if the limit (4.2) does not hold for some first-order ϕ then (4.6) does not hold – i.e. Duplicator doesn't almost always win. Suppose some first-order ϕ does not have a zero-one law. Then there would exist an $\epsilon > 0$ such that for all N, there exists an $m > N$ such that $\Pr[\mathcal{G}_m \models \phi] > \epsilon$ (otherwise the zero-one law could hold) and similarly, there exists an $n > N$ such that $\Pr[\mathcal{G}_n \models \phi] < 1 - \epsilon$. Using Theorem 3.30, let k be such that Spoiler wins $\texttt{EHR}[\mathcal{G}_n, \mathcal{G}_m, k]$ whenever $\mathcal{G}_n \models \phi$ but $\mathcal{G}_m \not\models \phi$. Since $\mathcal{G}_n$ and $\mathcal{G}_m$ are independent events we have

$$\Pr[\mathcal{G}_m \models \phi \wedge \mathcal{G}_n \not\models \phi] = \Pr[\mathcal{G}_m \models \phi] \Pr[\mathcal{G}_n \not\models \phi] > \epsilon(1 - (1-\epsilon)) = \epsilon^2 > 0.$$

We then have that Spoiler wins with nonzero probability, thus (4.6) does not hold.

Now we'll prove that if (4.2) *does* hold for every first-order ϕ then (4.6) also holds. Fix some k. For each $\alpha \in \texttt{EHRC}[k]$, let A_α be the characteristic first-order property of α. Set $\epsilon_\alpha = \lim_{n\to\infty} \Pr[\mathcal{G}_n \models A_\alpha]$. By assumption, ϵ_α is always either zero or one. For $\alpha \neq \beta$, the events $\mathcal{G}_n \models A_\alpha$ and $\mathcal{G}_n \models A_\beta$ are disjoint (since they correspond to disjoint parts of the partition $\texttt{EHRC}[k]$ of graphs). It follows that $\epsilon_\alpha = \epsilon_\beta = 1$ is *not* possible. In other words, at most one ϵ_α can be equal to one. Moreover, because these events partition the probability space,

$$\sum_{\alpha \in \texttt{EHRC}[k]} \Pr[\mathcal{G}_n \models A_\alpha] = 1. \tag{4.7}$$

Taking the limit of (4.7) we get $\sum_{\alpha \in \texttt{EHRC}[k]} \epsilon_\alpha = 1$. Thus at least one $\epsilon_\alpha = 1$. It follows that there is a unique $\epsilon_\alpha = 1$.

Let $\epsilon > 0$. By assumption, there exists an $N > 0$ such that if $n > N$, then $\mathcal{G}_n \models A_\alpha$ with probability greater than $1 - \epsilon$. Since $\mathcal{G}_n$ and $\mathcal{G}_m$ are independently chosen, for $n, m > N$, the probability that both $\mathcal{G}_n \models A_\alpha$ and $\mathcal{G}_m \models A_\alpha$ is greater than $(1-\epsilon)(1-\epsilon) = (1-\epsilon)^2$. By definition, when $\mathcal{G}_n$ and $\mathcal{G}_m$ are in the same $\equiv_k$ class Duplicator wins $\texttt{EHR}[\mathcal{G}_n, \mathcal{G}_m, k]$. The ϵ is arbitrary so the limit in (4.6) follows. □

We'll now give a quick preview of these ideas applied in contexts other than graphs before returning to one more proof of the zero-one law.

5. Ehrenfeucht Game for non-Graph Structures

We defined first-order formulae "on graphs." The thing that tied our definition to graphs as opposed to any other structure was that the only available symbols other than basic logical grammar are the relations $=$ and $\sim$. We could allow different relations like for instance $<$, but that would only make sense if we're dealing with structures that have an order to them. Formally, we defined first-order formulae "in the language of graphs." We could do so in other "languages" and thus find similarly powerful tools in contexts outside of graph theory.

Definition 5.1. A *language* $\mathcal{L}$ is a collection of symbols of three distinguished kinds: "function symbols," "relation symbols," and "constant symbols." To each function symbol f and each relation symbol R, there is associated a number n_f and n_R respectively indicating *arity*, meaning the number of arguments to be taken.

Indeed this definition seems vapidly general. That is okay. All this definition does for us is establish what the valid words are – not how to use them nor what they mean. Meaning comes in when considering an "$\mathcal{L}$-structure."

Definition 5.2. An $\mathcal{L}$-*structure* $\mathcal{M}$ is a set M together with "interpretations" of the symbols in $\mathcal{L}$:

- any function symbol $f \in \mathcal{L}$ associated to n_f is given the meaning of a specific function $f^{\mathcal{M}} : M^{n_f} \to M$
- any relation symbol $R \in \mathcal{L}$ associated to n_R is given the meaning of a specific relation $R^{\mathcal{M}} \subseteq M^{n_R}$
- any constant symbol $c \in \mathcal{L}$ is given the meaning of a specific distinguished element of M.

Examples clarify these clunky definitions. Equality $=$ is generally taken to be built into every language, but we will point it out explicitly as an available relation.

Example 5.3 (Graphs)**.** We call $\mathcal{L}_{\text{Graph}} \coloneqq \{\sim, =\}$ where $\sim$ and $=$ are binary (arity of 2) relation symbols, the *language of graphs*. A graph happens to be an $\mathcal{L}_{\text{Graph}}$-structure where M is the vertex set, $\sim$ is interpreted as adjacency, and $=$ is interpreted as equality. However, $\mathcal{M}$ being an $\mathcal{L}_{\text{Graph}}$-structure is not enough for $\mathcal{M}$ to satisfy our earlier definition of a graph: an $\mathcal{L}_{\text{Graph}}$-structure could have a vertex adjacent to itself. The next thing we'd need is the *theory of graphs*: a set of first-order axioms for graphs.

Recall the discussions of free variables in Definition 3.28. A *sentence* is a well-formed logical formula that has no free variables.

Definition 5.4. A set of sentences is called a *theory*. A set of *axioms* of a theory T is a subset $\Sigma \subseteq T$ such that any model of Σ is also a model of T.

This notion of axioms may be different than what one has in mind. If that is the case, one likely considers axioms to be a (small) subset of T so that any theorem that can be derived from T can also be derived from Σ. These definitions turn out to be equivalent in the first-order world, but we address logical inference in the next section where we actually use it in a proof.

Exercise 5.5. Translate the axioms for graphs in Definition 2.2 into first-order sentences in the language of graphs.

Here are some examples of new languages and structures.

Example 5.6. Consider the language of groups: $\mathcal{L}_{\text{Group}} \coloneqq \{+, =, 0\}$. We see our first function and constant symbols: $+$ is a function symbol taking two variables and 0 is a distinguished element, which would represent the group identity.

Exercise 5.7. Formally state the standard axioms of the theory of groups.

Example 5.8 (Linear Orderings)**.** We call $\mathcal{L}_{\text{LO}} \coloneqq \{<, =\}$ – both binary relation symbols – the *language of linear orderings.* We'll stop formally defining our interpretation of the symbols when it is clear from experience what they should mean, but the distinction between symbols and their meaning remains important. An example of an $\mathcal{L}_{\text{LO}}$-structure is the reals: $(\mathbb{R}, <, =)$.

The natural question is whether the Ehrenfeucht game generalizes to other structures. Indeed it does: the general Ehrenfeucht game goes the same way as before, except now it's not just an isomorphism with respect to $\sim$ and $=$, but with respect to all relations, functions, and constants of the language in question. We'll see this in action in Proposition 5.12.

The general version of Theorem 3.22 is as follows:

Theorem 5.9. *Let $\mathcal{L}$ be a finite language and $\mathcal{M}, \mathcal{N}$ be $\mathcal{L}$-structures.*

(1) For all $\equiv_k$ classes α, there is a characteristic first-order sentence ϕ of quantifier depth k such that $\mathcal{M} \models \phi \iff \mathcal{M} \in \alpha$.
(2) $\mathcal{M} \equiv_k \mathcal{N}$ iff for any first-order sentence ϕ of quantifier depth k, $\mathcal{M} \models \phi \iff \mathcal{N} \models \phi$.

Proof. See [5]. Section 2.4 of [6] is also helpful. □

The proof is more or less the same, just handling arbitrary relations, functions, and constants instead of only two binary relations $(\sim, =)$.

Things get more interesting when we consider a particular theory in a language $\mathcal{L}$ which not all $\mathcal{L}$-structures model. Let's take the theory of dense linear orderings without endpoints.

Definition 5.10. The theory of *dense linear orderings* (without endpoints) T_{DLO} is axiomatized by the regular properties of linear orderings: $<$ being transitive, the property $(\forall x)(\forall y)(x < y \lor x = y \lor x > y)$, etc. together with $(\forall x)(\exists y)(x < y)$, $(\forall x)(\exists y)(y < x)$, and the property of density: $\forall x, \forall y, x < y \implies \exists z : x < z < y$.

Remark 5.11. Though by now the reader probably has some experience with analysis, it is worth pointing out that the integers $(\mathbb{Z}, <, =)$ are an $\mathcal{L}_{\text{LO}}$-structure, but are not dense.

If for every k, we have $\mathcal{M} \equiv_k \mathcal{N}$, then we say $\mathcal{M}$ and $\mathcal{N}$ are *elementarily equivalent* and simply write $\mathcal{M} \equiv \mathcal{N}$. This is another way of saying that as far as one can probe with first-order logic, $\mathcal{M}$ and $\mathcal{N}$ look like the same structure. Now let's play the Ehrenfeucht game on $(\mathbb{R}, <, =)$ and $(\mathbb{Q}, <, =)$ and see what falls out.

Proposition 5.12. *The linear orderings $(\mathbb{R}, <, =)$ and $(\mathbb{Q}, <, =)$ are elementarily equivalent.*

Proof. Fix k. Suppose the positions are equivalent and without loss of generality, Spoiler picks $i_{\mathbb{R}}$. If $i_{\mathbb{R}} = j_{\mathbb{R}}$ for some $j < i$ then Duplicator just picks $j_{\mathbb{Q}}$ and by assumption, the positions remain equivalent. If $i_{\mathbb{R}}$ is a new vertex, let $G \coloneqq \{1 \le j < i : j_{\mathbb{R}} > i_{\mathbb{R}}\}$ and $L \coloneqq \{1 \le j < i : j_{\mathbb{R}} < i_{\mathbb{R}}\}$. Let $g \coloneqq \min_{j \in G} j_{\mathbb{Q}}$ and let $l \coloneqq \max_{j \in L} j_{\mathbb{Q}}$. Note that $l < g$. By density, take y such that $l < y < g$. Duplicator can pick y. We have shown for arbitrary k, $(\mathbb{R}, <, =) \equiv_k (\mathbb{Q}, <, =)$ and thus we are done. □

The way in which density works in this proof is analogous to the $(k-1)$-Alice's Restaurant property in Example 3.15. We'll see another similar argument in the next section.

Remark 5.13. Note that when we say $i_{\mathbb{R}} < j_{\mathbb{R}}$ we do *not* mean $i < j$, but rather that the real number marked by $i_{\mathbb{R}}$ is less than the real number marked by $j_{\mathbb{R}}$.

This proposition points out importantly that elementary equivalence is *not* isomorphism, though isomorphism implies elementary equivalence. More importantly, it shows that the difference between $\mathbb{Q}$ and $\mathbb{R}$ cannot be expressed in a first-order formula. In particular, it has the following fun corollary.

Corollary 5.14. *The least upper bound property is not first order.*

The theory of dense linear orderings is a surprisingly interesting object of study. However, theories are not only interesting as static objects of study, but also as dynamic tools in proofs. We demonstrate this with another proof of the zero-one law for graphs in the following section.

6. Another Logical Proof of the Zero-One Law

We present another proof that has some quite beautiful features, but we will have to accept a couple of powerful standard theorems from logic on faith. If that is unsettling as it was originally to the author, hopefully the first proof in Section 4 was convincing enough and some ideas from this proof can still be appreciated.

Remark 6.1. If the reader has some exposure to logic, different sources sometimes use different names or slightly different versions of the big theorems from logic that we will use. We follow the conventions of [7].

We briefly remarked on the notion of logical inference in the previous section. Logical inference is the application of a finite set of rules for going from previous logical expressions to new ones. The rules are typically taken to be "modus ponens," i.e. if $(P \implies Q) \wedge P$ then Q, as well as a couple more equally reasonable rules.

However, there are only a few features of proof that we need to accept, and they should be pretty reasonable for a mathematics student:

(1) A proof is a finite list of logical expressions,
(2) and those expressions are linked by the application of said rules to some combination of a) axioms and/or b) previously proved statements.

This discussion of formal proofs is fairly informal. For a formal treatment, any standard text on logic should do.

Notations 6.2. If T is a theory and ϕ is a first-order sentence that can be derived given the rules of logical inference, we write $T \vdash \phi$.

Recall that if a property ϕ is true of a graph G, we write $G \models \phi$ and say G models ϕ. If G is a graph that models every sentence in a theory T then we write $G \models T$ and say G is a model of T. One may reasonably ask the question, if $G \models T$ and I write a proof that $T \vdash \phi$, then does G model ϕ? The answer in the first-order world is yes! It is provided by Gödel's Completeness Theorem.

Theorem 6.3 (Incomplete Version of Gödel's Completeness Theorem)**.** *Let T be a first-order theory. Let ϕ be a sentence in the same (countable) language. The following are equivalent:*

- $T \vdash \phi$
- *every countable model of T satisfies ϕ.*

We'll use an even more incomplete version of Gödel's completeness theorem. As one might think, a consistent theory is one in which contradictory formulae are not derivable.

Theorem 6.4 (Our Version of Gödel's Completeness Theorem)**.** *Any consistent theory has a countable model.*

Proof. We prove the contrapositive. Suppose T has no countable model. Then it is vacuously true that any countable model of T satisfies $\phi \wedge \neg\phi$. By Theorem 6.3, $T \vdash (\phi \wedge \neg\phi)$, i.e T is inconsistent. □

We'll need one more big theorem from logic.

Theorem 6.5 (Our Version of Lowenheim-Skolem Theorem)**.** *If T is a theory with no finite models and a unique countable model up to isomorphism then T is complete.*

Wait! What is "completeness?!"

Definition 6.6. A theory T is *complete* if for any first-order sentence ϕ, either $T \vdash \phi$ or $T \vdash \neg\phi$.

What does a complete theory look like? The theory of dense linear orders without endpoints that we saw in the last section happens to be complete. We'll see what a completeness proof can look like during our new proof of the zero-one law.

Quick Outline. We'll construct a clever theory T such that we can apply the Lowenheim-Skolem theorem to determine T is complete. We'll then use the completeness of T to address arbitrary first-order properties and apply features of T as well as the fact that proof is finite to derive the result. □

We'll use the following fact.

Proposition 6.7. *If there exists a countable graph that satisfies the Alice's Restaurant property, it must be unique up to isomorphism.*

Proof. Suppose R_1 and R_2 have the Alice's Restaurant property. Since they are both countable, enumerate the vertices of R_1 as $x_1, x_2, x_3, ...$ and enumerate the vertices of R_2 as $y_1, y_2, y_3...$ The "back-and-forth" idea we will use is a common idea in logic. We will construct an isomorphism f between R_1 and R_2. First we will map $x_1 \mapsto y_1$. We will inductively define the isomorphism in two steps. On the first step we will look at the least i such that x_i has not been assigned a y in the range. Consider the neighbors N_i of x_i and the non-neighbors N_i^c of x_i that have been assigned elements in the range of the map. By the Alice's Restaurant Property, R_2 has a y such that $y \sim f(N_i)$ and $y \not\sim f(N_i^c)$. Assign $x_i \mapsto y$. Next we can similarly select the least j such that no $x \in R_1$ maps to y_j. Follow the same procedure and by the Alice's Restaurant Property, there exists an $x \in R_1$ such that $x \mapsto y_j$ keeps f being an isomorphism. Rinse and repeat ad infinitum. We get an isomorphism in the limit. □

Second Proof of Zero-One Law. Let T be the theory consisting of all $A_{r,s}$ extension properties. Recall from Proposition 4.3 that the $A_{r,s}$ extension properties all hold almost surely. Suppose $T \vdash \phi$. Since proof is finite, only finitely many sentences from T are used in the proof. Call them $A_1, ..., A_k$. By the completeness theorem, a model of $A_1 \wedge \cdots \wedge A_k$ is a model of ϕ. So if $\mathcal{G}_n \not\models \phi$ it follows that for some i, $\mathcal{G}_n \models \neg A_i$. Since each A_i holds almost surely, we have by the union bound

$$\Pr[\mathcal{G}_n \models \neg A_1 \vee \cdots \vee \mathcal{G}_n \models \neg A_k] \leq \Pr[\mathcal{G}_n \models \neg A_1] + \cdots + \Pr[\mathcal{G}_n \models \neg A_k] \to 0$$

as $n \to \infty$. What have we just learned? If $T \vdash \phi$ then ϕ must hold almost surely. Moreover, since ϕ holds almost surely, $\neg\phi$ holds almost never. The big takeaway is that closing T under logical inference we get a consistent theory. Going forward, by T we mean the closure of T under logical inference. By our version of Gödel's Completeness theorem, since T is consistent, it has a countable model. There exists a model for T, which must satisfy the Alice's Restaurant property so by Proposition 6.7, T must have an infinite unique countable model up to isomorphism. Call that model R.

With R, we can apply Lowenheim-Skolem to deduce that T must be complete. This is huge! We can now address *any* first-order ϕ and we're near the end of our proof. Let ϕ be a first-order property. Since T is complete, either $T \vdash \phi$ or $T \vdash \neg\phi$. If $T \vdash \phi$ then as we saw earlier in this proof, ϕ must hold almost surely. Similarly, $T \vdash \neg\phi$ implies $\neg\phi$ holds almost surely. This applies for any first-order ϕ so we have proved the zero-one law. □

A couple of remarks. It should be pointed out that the structure R that we use is actually called the "Rado graph." This is a highly symmetric structure with many interesting properties. For a survey on the topic, read Cameron's chapter (the first one) in "The Mathematics of Paul Erdős" ([2]). One such property is the following.

Exercise 6.8. Show that with probability 1, a countably infinite random graph $\mathcal{G} \sim G(\aleph_0, 1/2)$ satisfies the Alice's Restaurant Property. Corollary: with probability 1, a random graph on countably many vertices is isomorphic to the Rado graph R.

Erdős and Spencer had a fun comment: Exercise 6.8 "demolishes the theory of infinite random graphs" ([2]). Indeed, there is such a thing as *too much* structure. Yet, as Spencer points out in [7], it is quite appealing how this proof uses an infinite structure (R) to prove a statement that is exclusively about finite structures. That feature makes this proof particularly worthwhile, though it lacks the conceptual gratification of the Ehrenfeucht proof.

Acknowledgments

I would like to thank my mentor, Marc de Fontnouvelle. Logic would not have even been on my radar had you not been my mentor. Thank you for your patience, help, and insights.

Huge thanks are also in order for Professor May who organized the REU. This was an incredible learning experience I would not have had anywhere else. I would also like to thank Professor Rudenko for widening my perspective in mathematics and Professor Babai for helping me improve leaps and bounds at problem solving – his problem sessions made a huge impression on me. Lastly, I would like to thank my peers for making a great environment for learning and collaboration.

References

[1] Rabah Amir and Igor V. Evstigneev. On Zermelo's Theorem. Journal of Dynamics and Games. 2017.

[2] Peter J. Cameron. The Random Graph.

[3] Reinhard Diestel. Graph Theory. Fifth Edition. Springer-Verlag Berlin Heidelberg. 2017.
[4] Thomas Kern. https://trkern.github.io/efg.html.
[5] Asa Levi. The Ehrenfeucht Game and First Order Logic. 2006.
[6] David Marker. Model Theory: An Introduction. Springer. New York, New York. 2002.
[7] Joel Spencer. The Strange Logic of Random Graphs. Springer Publishing Company, Incorporated. 2010.
[8] Terence Tao. https://terrytao.wordpress.com/tag/godel-completeness-theorem/
[9] William Weiss and Cherie D'Mello. Fundamentals of Model Theory.