

MULTI-STRUCTURAL GAMES AND NUMBER OF QUANTIFIERS*

RONALD FAGIN ^a, JONATHAN LENCHNER ^b, KENNETH W. REGAN ^c,
AND NIKHIL VYAS ^{d,†}

^a IBM Research - Almaden, San Jose, CA
e-mail address: fagin@us.ibm.com

^b IBM T.J. Watson Research Center, Yorktown Heights, NY
e-mail address: lenchner@us.ibm.com

^c Department of CSE, University at Buffalo, Amherst, NY
e-mail address: regan@buffalo.edu

^d Department of EECS, MIT, Cambridge, MA
e-mail address: nikhil@g.harvard.edu

ABSTRACT. We study multi-structural games, played on two sets \mathcal{A} and \mathcal{B} of structures. These games generalize Ehrenfeucht-Fraïssé games. Whereas Ehrenfeucht-Fraïssé games capture the quantifier rank of a first-order sentence, multi-structural games capture the number of quantifiers, in the sense that Spoiler wins the r -round game if and only if there is a first-order sentence ϕ with at most r quantifiers, where every structure in \mathcal{A} satisfies ϕ and no structure in \mathcal{B} satisfies ϕ . We use these games to give a complete characterization of the number of quantifiers required to distinguish linear orders of different sizes and we develop machinery for analyzing structures beyond linear orders.

1. INTRODUCTION

Model theory has a number of techniques for proving inexpressibility results. However, as noted in [Fag93], almost none of the key theorems and tools of model theory, such as the compactness theorem and the Löwenheim-Skolem theorems, apply to finite structures. Among the few tools of model theory that yield inexpressibility results for finite structures are Ehrenfeucht-Fraïssé games [Ehr61, Fra54], henceforth E-F games. In this paper we shall entirely be concerned with *finite* structures over *finite* relational vocabularies. For definitions of these concepts, we refer the reader to [Imm99].

The standard E-F game is played by “Spoiler” and “Duplicator” on a pair (A, B) of structures over the same first-order vocabulary τ , for a specified number r of *rounds*. In each round, Spoiler chooses an element from A or from B , and Duplicator replies by choosing

*This paper is an expanded and reorganized version of the LICS (Logic in Computer Science) 2021 paper [FLRV21a], which points to the Computer Science arXiv, version 1, entry [FLRV21b] for full proofs of its theorems.

[†]Supported by NSF CCF-1909429. Work Partially done while visiting IBM Research - Almaden.

an element from the other structure. In this way, they determine sequences of elements $a_1, \dots, a_r \in A$ and $b_1, \dots, b_r \in B$, repetitions allowed, which, in addition to any possible constants defined on A and B , define substructures A' of A and B' of B . The analysis can be phrased as attempting to define a sequence of functions f , g , and $f \cup g$. If any of these functions are not well-defined then Spoiler wins. The function $f: \{a_1, \dots, a_r\} \rightarrow \{b_1, \dots, b_r\}$ is defined by $f(a_i) = b_i$ for $i = 1, \dots, r$. If there are i and j with $1 \leq i < j \leq r$ such that $a_i = a_j$ but $b_i \neq b_j$, then f is not well-defined. The partial function $g: A \rightarrow B$ is defined so that $g(a) = b$ if a and b are associated with the same constant symbol in τ . If after a round of play there are two constant symbols c_i, c_j with $a \in A$ associated with both c_i and c_j , but different elements of B associated with c_i and c_j , then g is not well-defined. The joint partial function $f \cup g$ stipulates that $(f \cup g)(a) = f(a)$ if f is defined on a and $(f \cup g)(a) = g(a)$ if g is defined on a . If f and g are well-defined and agree on common elements, then $f \cup g$ is well-defined. If it is not well-defined, or if $f \cup g$ is not an isomorphism from its domain A' to its image B' , then Spoiler wins. Finally, if $f \cup g$ is well-defined, and $f \cup g$ is an isomorphism from A' to B' , then Duplicator wins. In this latter case of a Duplicator win, we say that the sequences $\langle a_1, \dots, a_r \rangle$ and $\langle b_1, \dots, b_r \rangle$ of selected elements “give rise to a partial isomorphism” between A and B .

The equivalence theorem for E-F games [Ehr61, Fra54] characterizes the minimum *quantifier rank* of a sentence ϕ over τ that is true for A but false for B . The quantifier rank $\text{qr}(\phi)$ is defined as zero for a quantifier-free sentence ϕ , and inductively:

$$\begin{aligned} \text{qr}(\neg\phi) &= \text{qr}(\phi), \\ \text{qr}(\phi \vee \psi) &= \text{qr}(\phi \wedge \psi) = \max\{\text{qr}(\phi), \text{qr}(\psi)\}, \\ \text{qr}(\forall x\phi(x)) &= \text{qr}(\exists x\phi(x)) = \text{qr}(\phi) + 1, \end{aligned}$$

Theorem 1.1 (Equivalence Theorem for E-F Games). *Spoiler wins the r -round E-F game on (A, B) if and only if there is a 1st order sentence ϕ of quantifier rank at most r such that $A \models \phi$ while $B \models \neg\phi$.*

The “if” direction of this theorem is fairly easy to prove by induction on r . This is the “useful” direction, which is used to prove inexpressibility results. The “only if” direction is somewhat tricky to prove; intuitively, it tells us that any technique for proving that a certain property cannot be defined by a first-order sentence with a certain quantifier rank can, in principle, be replaced by a proof via E-F games. See [Imm99, Lib12] for a proof and extended discussion.

We investigate a variant of E-F games that we call *multi-structural games*. These games make Duplicator more powerful and characterize the *number* of quantifiers rather than *quantifier rank*. It is straightforward to see that the minimum number of quantifiers needed to define a property P is the same as the minimum size of the quantifier prefix of a sentence in prenex normal form that is needed to define property P . This is because converting a sentence into prenex normal form does not increase the number of quantifiers.

As we discovered during review of this paper’s conference version [FLRV21a] and acknowledged there, an equivalent of our multi-structural game was described in the journal version of Neil Immerman’s paper “Number of Quantifiers Is Better Than Number of Tape Cells” [Imm81]. Its conference version [Imm79] did not mention the game.¹ In [Imm81],

¹We reference two personal communications with Immerman as [Imm21], one beforehand where he raised the related size game described in [AI03] and discussed in section 1.1, and one after our discovery that informed the present discussion.

Immerman called it the “separability game” and showed that it characterizes the number of quantifiers, without providing further results. Just prior to the conclusion of [Imm81], Immerman remarked,

“Little is known about how to play the separability game. We leave it here as a jumping off point for further research. We urge others to study it, hoping that the separability game may become a viable tool for ascertaining some of the lower bounds which are ‘well believed’ but have so far escaped proof.”

Indeed, as our paper shows, analysis of the multi-structural games is often quite confounding, with many delicate issues.

We now define the rules of the multi-structural game (henceforth, “MS game”). There are again two players, Spoiler and Duplicator, and there is a fixed number r of rounds. Instead of being played on a pair (A, B) of structures with the same vocabulary (as in an E-F game), it is played on a pair $(\mathcal{A}, \mathcal{B})$ of sets of structures, all with the same vocabulary. For k with $0 \leq k \leq r$, by a *labeled structure* after k rounds, we mean a structure along with a labeling of which elements were selected from it in each of the first k rounds. Let $\mathcal{A}_0 = \mathcal{A}$ and $\mathcal{B}_0 = \mathcal{B}$. Thus, \mathcal{A}_0 represents the labeled structures from \mathcal{A} after 0 rounds, and similarly for \mathcal{B}_0 . If $1 \leq k < r$, let \mathcal{A}_k be the labeled structures originating from \mathcal{A} after k rounds, and similarly for \mathcal{B}_k . In round $k + 1$, Spoiler either chooses an element from each member of \mathcal{A}_k , thereby creating \mathcal{A}_{k+1} , or chooses an element from each member of \mathcal{B}_k , thereby creating \mathcal{B}_{k+1} . Duplicator responds as follows. Suppose that Spoiler chose an element from each member of \mathcal{A}_k , thereby creating \mathcal{A}_{k+1} . Duplicator can then make multiple copies of each labeled structure of \mathcal{B}_k , and choose an element from each copy, thereby creating \mathcal{B}_{k+1} . Similarly, if Spoiler chose an element from each member of \mathcal{B}_k , thereby creating \mathcal{B}_{k+1} , Duplicator can then make multiple copies of each labeled structure of \mathcal{A}_k , and choose an element from each copy, thereby creating \mathcal{A}_{k+1} . Duplicator wins if there is some labeled A in \mathcal{A}_r and some labeled B in \mathcal{B}_r where the labelings (in addition to any constants) give rise to a partial isomorphism in the same sense as in an E-F game. Otherwise, Spoiler wins.

Note that on each of Duplicator’s moves, Duplicator can make “every possible choice,” via the multiple copies. Making every possible choice creates what we call the *oblivious strategy*. It is easy to see that Duplicator has a winning strategy if and only if the oblivious strategy is a winning strategy.

We shall prove the following theorem. It is analogous to Theorem 1.1 for ordinary E-F games.

Theorem 1.2. *Equivalence Theorem for MS Games:* *Spoiler wins the r -round MS game on $(\mathcal{A}, \mathcal{B})$ if and only if there is a 1st order sentence ϕ with at most r quantifiers such that $A \models \phi$ for every $A \in \mathcal{A}$ while $B \models \neg\phi$ for every $B \in \mathcal{B}$.*

We now give an interesting refinement of the Equivalence Theorem (although, as we shall discuss, it does not seem to directly imply the Equivalence Theorem). Let $Q_1 \cdots Q_r$ be a sequence of quantifiers. We now define the “ $Q_1 \cdots Q_r$ MS game”. It is an r -round MS game, with the following restrictions on Spoiler. If Q_k is an existential quantifier, then Spoiler’s k th move must be in \mathcal{A} , and otherwise it must be in \mathcal{B} . We then have the following result.

Theorem 1.3. *Fixed Prefix Equivalence Theorem for MS Games:* *Spoiler wins the $Q_1 \cdots Q_r$ MS game on $(\mathcal{A}, \mathcal{B})$ if and only if there is a 1st order sentence ϕ in prenex normal form with exactly r quantifiers, in the order $Q_1 \cdots Q_r$, such that $A \models \phi$ for every $A \in \mathcal{A}$ while $B \models \neg\phi$ for every $B \in \mathcal{B}$.*

On the face of it, Theorem 1.3 does not seem to directly imply Theorem 1.2, for the following reason. It is a priori conceivable that Spoiler’s winning strategy in an r -round game is to move first in \mathcal{A} , and then, depending on Duplicator’s response, to move in either \mathcal{A} or \mathcal{B} . So there is then no prefix $Q_1 \cdots Q_r$ dictating where Spoiler must move. In fact, in an E-F game (but not in a MS game²), it can indeed happen that Spoiler can win in 2 rounds, and where Spoiler plays in the second round depends on how Duplicator played in the first round. An example of where this phenomenon can take place in a 2-round E-F game is via the sentence $\exists x(\forall yB(x, y) \wedge \exists yR(x, y))$. The proof of Theorem 1.2 appears in Section 2. The proof of Theorem 1.3 is almost the same as the proof of Theorem 1.2 and requires just very minor changes.

There is an interesting and non-obvious difference between E-F games and MS games. Let us say that a player makes a move “on top of” a previous move if the player selects an element c of a structure, and the same element c had been selected by either player in an earlier round. It is easy to see that in an E-F game, it never helps Spoiler to make a move on top of a previous move (it only wastes a round). On the other hand, in MS games the issue of playing on top of a previous move is a frequent consideration for us. In fact, a detailed analysis shows that playing a move on top of a previous move may be a necessary part of a winning Spoiler strategy (see Observation B.1).

We now give an example (Figure 1) that shows differences between the E-F game and the MS game, and what they say about quantifier rank vs. number of quantifiers. Consider the following two structures B (for “Big”) and L (for “Little”), over $\tau = \{<\}$, where $<$ is the binary “less than” relation. The vertex labels are not part of the structures. Elements that appear to the left, within the same structure, are considered to be less than elements to the right. B is a linear order on 3 elements and L is a linear order on two elements. In the text of this paper we write $B(i)$ (or $L(i)$) to denote the i th element in the linear order B (respectively, L), while in the figures, for economy of space, we label the i th vertex instead by Bi (respectively, Li). Further, rather than use the notation $<(x, y)$ we shall use the



Figure 1: An example showing the difference between MS and E-F games.

customary $x < y$ notation.

Suppose first that the number r of rounds is 2. We show that Spoiler wins the E-F game. On Spoiler’s first move, Spoiler selects vertex $B(2)$ in B . Duplicator must select either $L(1)$ or $L(2)$ in L . If Duplicator chooses $L(1)$, then Spoiler selects $B(1)$ in B . After Duplicator selects $L(2)$ in L Spoiler wins since the mapping given by $B(2) \mapsto L(1)$ and $B(1) \mapsto L(2)$ fails to be a partial isomorphism because the “less than” relationship is flipped. If Duplicator had instead selected $L(2)$ in the first round, then Spoiler would have won, by a similar argument, by selecting $B(3)$ in B in the second round.

The fact that Spoiler wins the 2-round game over (B, L) tells us (by Theorem 1.1) that there is a sentence ϕ of quantifier rank at most 2 such that $B \models \phi$ while $L \models \neg\phi$. Such a sentence is: $\exists x(\exists y(y < x) \wedge \exists y(x < y))$.

²Since we can assume Duplicator plays obliviously, Spoiler’s moves are not conditioned on what Duplicator plays.

Now let us consider the 2-round MS game over $(\mathcal{B}, \mathcal{L})$, where $\mathcal{B} = \{B\}$ and $\mathcal{L} = \{L\}$. We show that unlike the 2-round E-F game, in the 2-round MS game, Duplicator wins. It is easy to see that Duplicator wins if Spoiler's first-round move is anything other than $B(2)$. Let us see what happens if Spoiler's first move is $B(2)$, which was a winning move for Spoiler in the 2-round E-F game. Then on Duplicator's first move, Duplicator makes a second copy of L and in one copy, call it L_1 , Duplicator selects $L(1)$, and in the other copy, call it L_2 , Duplicator selects $L(2)$. Let us now consider Spoiler's possible second round responses. Suppose first that Spoiler's second round move is in B . If Spoiler selects $B(3)$, then Duplicator selects $L(2)$ in L_1 and the mapping $B \rightarrow L_1$ such that $B(2) \mapsto L(1), B(3) \mapsto L(2)$ yields a partial isomorphism. On the other hand, if Spoiler selects $B(1)$, then Duplicator selects $L(1)$ in L_2 and $B \rightarrow L_2$ such that $B(1) \mapsto L(1), B(2) \mapsto L(2)$ yields a partial isomorphism. Section 4 will complete the analysis of a Duplicator win. Since Duplicator wins the 2-round game, it follows by Theorem 1.2 there is no sentence with just two quantifiers that distinguishes \mathcal{B} from \mathcal{L} .

The focus of our analysis of MS games in this paper is finite linear orders. Henceforth, all linear orders are assumed to be finite. In the case of E-F games, one has the following:

Theorem 1.4 [Ros82]. *Let $f(r) = 2^r - 1$. In an r -round E-F game played on two linear orders of different sizes, Duplicator wins if and only if the size of the smaller linear order is at least $f(r)$.*

Since part of the proof of Theorem 1.4 is left to an exercise in [Ros82], we give a proof in Appendix A. Further, the proof illustrates a simple recursive idea that is surprisingly not available to us in the analysis of linear orders from the vantage point of MS games.

Theorem 1.4 together with Theorem 1.1 imply that $f(r)$ is the maximum value k such that a sentence of quantifier rank r can distinguish linear orders of size k and above from those of size smaller than k .

In analogy to this function f , and in an effort to arrive at a parallel theorem to Theorem 1.4, we make the following definition.

Definition 1.5. Define the function $g : \mathbb{N} \rightarrow \mathbb{N}$ such that $g(r)$ is the maximum number k such that there is a sentence with r quantifiers that can distinguish linear orders of size k or larger, from linear orders of size less than k .

To see that g is well-defined, observe that the sentence

$$\exists x_1 \cdots \exists x_r \bigwedge_{1 \leq i < r} x_i < x_{i+1}, \quad (1.1)$$

distinguishes linear orders of size r or larger from linear orders of size less than r . Furthermore, there are only finitely many inequivalent sentences with up to r quantifiers that include only the relation symbols $<$ and $=$, some fraction of which distinguish linear orders of some size k or greater from linear orders of size less than k . There is therefore a maximum such $k \geq r$, which is then $g(r)$.

After building up quite a bit of machinery we eventually arrive at the following:

Theorem 1.6. *The function g takes on the following values: $g(1) = 1, g(2) = 2, g(3) = 4, g(4) = 10$, and for $r > 4$,*

$$g(r) = \begin{cases} 2g(r-1) & \text{if } r \text{ is even,} \\ 2g(r-1) + 1 & \text{if } r \text{ is odd.} \end{cases}$$

The value $g(3) = 4$ is a curious anomaly. If it had turned out that $g(3) = 5$, then the entire induction could be founded on $r = 1$. The proof of Theorem 1.6 is a careful mathematical journey to restore this induction founded instead at $r = 4$.

The following theorem for MS games is the analog of Theorem 1.4 for E-F games, and describes precisely when Duplicator (alternatively, Spoiler) wins r -round MS games on two linear orders of different sizes.

Theorem 1.7. *In an r -round MS game played on two linear orders of different sizes Duplicator has a winning strategy if and only if the size of the smaller linear order is at least $g(r)$.*

It is important to note that neither the *if* nor *only if* portion of this theorem is implied by the definition of g .

Both E-F games and MS games are used to prove inexpressibility results by showing there is a winning strategy for Duplicator. It is typically easier to demonstrate a winning strategy for Duplicator in E-F games than in MS games, for several reasons. First, it is easier to reason about only two structures at a time rather than about many structures at a time. Second, in MS games, there is a tactic available to Spoiler (that is of no use in E-F games) to make one move on top of an earlier move in one of the structures, and this can greatly complicate the analysis. On the other hand, in MS games, Duplicator has the advantage of being able to make multiple copies of structures and make different moves on the various copies. This feature can be very useful in proving a winning strategy for Duplicator.

A similar phenomenon of modifying the rules of the game to make it easier for Duplicator to win arose in defining and making use of Ajtai-Fagin games [AF90] rather than making use of the originally defined Fagin games [Fag75] for proving inexpressibility results in monadic existential second-order logic (called “monadic NP” in [FSV95]). In Fagin games, there is a coloring round (a choice of the combinations of existentially-quantified monadic predicates) where Spoiler colors A then Duplicator colors B , and then an ordinary E-F game is played on the colored structures. In Ajtai-Fagin games, Spoiler must commit to a coloring of A without knowing what the other structure B is. Fagin, Stockmeyer, and Vardi [FSV95] use Ajtai-Fagin games to give a much simpler proof that connectivity is not in monadic NP than Fagin’s original proof in [Fag75]. In extending MS games to second-order logic, which we think is an interesting and important future step (and which is straightforward to define), these games can easily simulate Ajtai-Fagin games. This is because we can replace structure A by the singleton set $\mathcal{A} = \{A\}$, and replace the structure B by a set \mathcal{B} that contains all possible choices for B that Duplicator might choose in the Ajtai-Fagin game after Spoiler colors A .

1.1. Related work. Since our results can be viewed as giving information about the size of prefixes of sentences in prenex normal form, we begin by discussing some other papers that focus on such prefixes.

Rosen [Ros05] shows that there is a strict prefix hierarchy, based on the prefixes of sentences written in prenex normal form. The proof involves standard E-F games.

Dawar and Sankaran [DS21] consider E-F games, each of which focuses on a fixed prenex prefix. For example, there is one game that deals with the prenex prefix $\exists\forall\exists$. For each of these prefixes, they define an E-F game on a pair (A, B) of structures. For example, in the $\exists\forall\exists$ game, Spoiler must move first in A , then in B , and then in A . Their Theorem 2.3 says that Spoiler has a winning strategy in a prefix game if and only if there is a sentence

in prenex normal form with exactly that prefix that is true about A but not about B . Unfortunately, the “only if” direction of their Theorem 2.3 is false [Daw21]. This is because if A is a linear order of size 5 and B is a linear order of size 4, and the prefix is $\exists\forall\exists$, then it turns out that Spoiler wins that 3-round prefix game, but it follows from our Theorem 1.6 that A and B agree on all sentences with at most three quantifiers, and in particular on all sentences $\exists x\forall y\exists z\phi(x, y, z)$, where ϕ is quantifier-free. Fortunately, in their paper, Dawar and Sankaran just make use of the “if” direction of Theorem 2.3, which is correct [Daw21].

We now discuss some papers that, like ours, modify E-F games by allowing a pair of sets of structures, rather than simply a pair of structures. Adler and Immerman [AI03] use a type of E-F game that involves a pair of sets of structures, where, as in our MS games, Duplicator can make multiple copies of structures and make different moves on them. Adler and Immerman’s concern is to obtain results about the size of sentences (rather than the number of quantifiers) in transitive closure logic (first-order logic with the transitive closure operator). The rules of the game are rather complicated, since it must deal with transitive closure logic and capture the size of sentences.

Hella and Vilander [HV19] build on Adler and Immerman’s game, and their goal is also to determine sentence size (but in modal logic). The rules of their game are also fairly complicated.

Grohe and Schweikardt [GS05] introduce a method (extended syntax trees) that corresponds to a game tree that is constructed by the two players in the Adler-Immerman game. They use these to study the size of sentences in the 2, 3 and 4-variable fragments of first-order logic on linear orders.

Lotfallah [Lot04] introduces a class of E-F-like games played on a pair of sets of structures rather than on a pair of single structures. In Lotfallah’s games, Duplicator cannot make multiple copies of structures. A follow-up paper by Lotfallah and Youssef [LY05] characterizes certain first and second order prefix types but does not involve sets of structures.

Hella and Väänänen [HV15], like Lotfallah, introduce a class of E-F-like games played on a pair of sets of structures rather than on a pair of single structures, where Duplicator cannot make multiple copies of structures. Hella and Väänänen use one variant of their game to characterize the size of sentences needed for separating sets of structures in propositional logic and a second variant to characterize the size of sentences needed for separating sets of structures in first-order logic. The first-order game is used for proving an exact bound on the size of existential sentences needed to define the length of linear orders.

1.2. Overview of the Sections. In Section 2 we prove the Equivalence Theorem 1.2. In Section 3 we establish certain preliminary terminology and notation and prove that the property of Duplicator having a winning strategy on two sets of structures gives rise to an equivalence relation over sets of structures with the same signature. The ensuing sections establish upper and lower bounds on the function $g(r)$ associated with Theorem 1.6 until we are able to observe that we have tight bounds in all cases. In Section 4 we establish upper bounds on $g(r)$, for 2 and 3. (The trivial tight bound $g(1) = 1$ is established in Section 3.) The natural next step would be to proceed to higher values of r using a type of recursive argument, but in Section 5 we show why the natural recursive argument for MS games does not work. For pedagogical reasons, in Section 6, we jump to establishing lower bounds for $g(r)$. We then jump back to establishing upper bounds in Section 7, where we introduce a new type of game, an MS game with “atoms”, which allows us to recurse and prove upper bounds for all r . The upper bounds are then seen to be tight with respect to our lower

bounds and hence, in Section 8, we are able to prove Theorem 1.6. The machinery that we have built up then enables a quick proof of Theorem 8.4, which is just the syntactic equivalent of Theorem 1.7 from the Introduction (there stated game theoretically). We note that while games with “atoms” are an important part of our upper bound proofs, the final sentences guaranteed by Definition 1.5 and Theorem 1.6 do not contain atoms. We present our conclusions in Section 9.

2. PROOF OF THE EQUIVALENCE THEOREM

To prove Theorem 1.2 we are going to add a special constant just prior to the play of each round that will help us maintain the induction. These constants, which we denote by c_1, \dots, c_r are in addition to whatever constants may exist in the vocabulary τ . We write $(A; \dots; c_1 \leftarrow a_1)$ to mean the structure obtained after assigning c_1 to the element a_1 in A , and so on.

Proof of Theorem 1.2. Both directions are proved by induction on the number r of rounds. For $r = 0$, Spoiler winning in 0 rounds means that for every $A \in \mathcal{A}$ and $B \in \mathcal{B}$, the restrictions A', B' of A and B (respectively) to their constants must be non-isomorphic. For every $A \in \mathcal{A}$, we can write a quantifier-free sentence ϕ_A that characterizes A' up to isomorphism, using equality to identify any coinciding constants. Then take ϕ to be the disjunction of all ϕ_A – note that even if \mathcal{A} is infinite, there are only finitely many distinct ϕ_A . Then $A \models \phi$ for all $A \in \mathcal{A}$. Now consider any $B \in \mathcal{B}$. We claim that $B \models \neg\phi$. As $\neg\phi$ is a conjunction, this is equivalent to $B \models \neg\phi_A$ for every A . Suppose not, then we would have $B \not\models \neg\phi_A$, i.e., $B \models \phi_A$. But because ϕ_A characterizes A' up to isomorphism, this would make B' isomorphic to A' , contradicting that Spoiler wins. Hence ϕ is a quantifier-free sentence that distinguishes \mathcal{A} and \mathcal{B} .

Conversely, if there is a quantifier-free sentence ϕ that distinguishes \mathcal{A} and \mathcal{B} , then there cannot exist $A \in \mathcal{A}$ and $B \in \mathcal{B}$ such that the restrictions A' and B' to the constants appearing in ϕ are isomorphic. Thus, Duplicator loses without further play.

Now suppose $r \geq 1$ and the equivalence is true for $r - 1$. We induct on the first move rather than the last move of the games. For the forward direction, suppose Spoiler can win—say by playing in \mathcal{B} . For each $B \in \mathcal{B}$, Spoiler selects an element $b \in B$. Duplicator replies by replicating every $A \in \mathcal{A}$ and playing every possible $a \in A$. Now let us assign the constant c_1 to the element played on each of the structures in \mathcal{A} and \mathcal{B} . The resulting game position $(\mathcal{A}^1, \mathcal{B}^1)$, which includes the new constant c_1 in each structure, is winnable in $r - 1$ rounds by Spoiler. By the induction hypothesis, there is a sentence ψ with $r - 1$ quantifiers that distinguishes \mathcal{B}^1 from \mathcal{A}^1 . Now define $\phi = (\exists x_1)\psi'$ where ψ' replaces all occurrences of c_1 in ψ by x_1 . (This is alright even in degenerate cases where c_1 does not occur in ψ .) For every $B \in \mathcal{B}$, we have $B \models \phi$ because there is a $b \in B$ such that $(B; \dots; c_1 \leftarrow b) \models \psi$, namely the b that Spoiler played in B . Hence it suffices to show that $A \models \neg\phi = (\forall x_1)\neg\psi'$ for every $A \in \mathcal{A}$. After Duplicator’s play, A was replaced by

$$\{(A; \dots; c_1 \leftarrow a_1), (A; \dots; c_1 \leftarrow a_2), \dots, (A; \dots; c_1 \leftarrow a_m)\},$$

where $A = \{a_1, \dots, a_m\}$. Since ψ distinguishes \mathcal{B}^1 from \mathcal{A}^1 , we have $(A; \dots, c_1 \leftarrow a_j) \models \neg\psi$ for each $j = 1, \dots, m$. It follows that $A \models (\forall x_1)\neg\psi'$. The case where Spoiler wins by playing in \mathcal{A} is handled symmetrically.

Going the other way, suppose ϕ is a prenex sentence with r quantifiers that distinguishes \mathcal{A} from \mathcal{B} . If the leading quantifier is \forall then $\neg\phi$ has leading quantifier \exists and distinguishes \mathcal{B}

from \mathcal{A} , so we can reason by symmetry. So let $\phi = (\exists x_1)\psi'$ for some ψ' and take ψ to be the sentence ψ' with x_1 replaced everywhere by the special constant symbol c_1 . For every $A \in \mathcal{A}$, $A \models \phi$, so there exists $a_1 \in A$ such that $(A; \dots; c_1 \leftarrow a_1) \models \psi$. Spoiler can play such an element a_1 in every A . Now every $B \in \mathcal{B}$ models $\neg\phi = (\forall x_1)\neg\psi'$. For every $b \in B$, Duplicator creates the structure $(B; \dots; c_1 \leftarrow b)$, but regardless of b , it models $\neg\psi$. Thus, ψ distinguishes the resulting set \mathcal{A}^1 from Duplicator's \mathcal{B}^1 , has $r - 1$ quantifiers, and includes c_1 along with any previous constants. By the induction hypothesis, Spoiler wins from $(\mathcal{A}^1, \mathcal{B}^1)$ in $r - 1$ rounds, so Spoiler wins from $(\mathcal{A}, \mathcal{B})$ in r rounds. \square

Proof of Theorem 1.3. In both directions of the induction, if the first quantifier Q_1 is fixed to be \exists or fixed to be \forall , then this limits which case can arise within the induction step, but does not affect its validity. \square

3. PRELIMINARIES

Definition 3.1. Let $\mathcal{A} = \{A\}$ and $\mathcal{B} = \{B\}$ be two *singleton sets of structures*. Write $\mathcal{A} \equiv_r \mathcal{B}$ iff Duplicator has a winning strategy for MS games of r rounds on \mathcal{A} and \mathcal{B} .

An important consequence of Theorem 1.2 is the following.

Lemma 3.2. *The relation \equiv_r is an equivalence relation between singleton sets of structures.*

Proof. That the relation \equiv_r is reflexive and symmetric follows immediately from the definition. For transitivity, suppose there are three singleton sets of structures, $\mathcal{A} = \{A\}$, $\mathcal{B} = \{B\}$ and $\mathcal{C} = \{C\}$ such that $\mathcal{A} \equiv_r \mathcal{B}$ and $\mathcal{B} \equiv_r \mathcal{C}$. By the Equivalence Theorem 1.2, $\mathcal{A} \equiv_r \mathcal{B}$ implies that \mathcal{A} and \mathcal{B} agree on the same sentences with at most r quantifiers. Similarly, $\mathcal{B} \equiv_r \mathcal{C}$ implies that \mathcal{B} and \mathcal{C} agree on the same set of sentences with at most r quantifiers. Hence, \mathcal{A} and \mathcal{C} agree on these same set of sentences. By the Equivalence Theorem again, it follows that $\mathcal{A} \equiv_r \mathcal{C}$. \square

Observation 3.3. Among non-singleton sets of structures, the property of there being a Duplicator-winning strategy between two sets does *not* induce an equivalence relation.

The induced relation between pairs of sets of structures can fail to be transitive. To see this, let A be a graph with no self-loops among the nodes, and let C be a graph with at least one node with a self-loop. Now let $\mathcal{A} = \{A\}$, $\mathcal{B} = \{A, C\}$, and $\mathcal{C} = \{C\}$. Duplicator wins a 1-round M-S game on \mathcal{A} and \mathcal{B} simply by focusing on A , and, analogously, Duplicator wins a 1-round M-S game on \mathcal{B} and \mathcal{C} simply by focusing on C . However, Spoiler wins the 1-round M-S game on \mathcal{A} and \mathcal{C} by playing on a node in C that has a self-loop.

We note that Definition 3.1 and Lemma 3.2 are incorrectly stated for sets of structures in both the LICS conference [FLRV21a] and extended ArXiV [FLRV21b] versions of this manuscript. The fallacious bit of reasoning is that in the case of non-singleton sets \mathcal{A}, \mathcal{B} of structures, the fact that Duplicator has a winning strategy in an r -round M-S game does *not* imply that all elements of \mathcal{A} and \mathcal{B} agree on every r -quantifier sentence, as was claimed – consider, for example, the sets \mathcal{A} and \mathcal{B} in the preceding paragraph. We thank Phokion Kolaitis for pointing out this error to us.

Before proceeding further let us establish some terminology that is intended to make the reading smoother. In cases where there are multiple linear orders on one side or another we often refer to the different linear orders as different “boards” that Spoiler and Duplicator play on.

When we say a “ K versus K' game” or a “ K vs. K' game”, we mean an MS game played on $(\mathcal{A}, \mathcal{B})$ where \mathcal{A} consists of a single linear order of size K , and \mathcal{B} consists of a single linear order of size K' .

We will typically play games where \mathcal{A} and \mathcal{B} each consist of a single linear order as above. In this context, as we did in the Introduction, we will use B to denote the *big* linear order and L to denote the *little* linear order.

As is standard in model theory, we assume a non-empty universe so all linear orders are of size at least 1. With reference to the function g in Definition 1.5, we begin by establishing:

Lemma 3.4. $g(1) = 1$.

Proof. The sentence $\exists x(x = x)$ is true for all linear orders so that $g(1) \geq 1$. Moreover, Duplicator can win 1-round MS games whenever both linear orders are of size 1 or greater, which implies that $g(1) \leq 1$. \square

4. TOWARDS ESTABLISHING UPPER BOUNDS ON $g(r)$

A potent tool for finding an upper bound k on the value of $g(r)$ will be to find strategies such that Duplicator can win r -round games on a *pair of singleton linear orders* whenever the sizes of the linear orders are at least k . All of our upper bounds are established in this manner.

Since we have established that $g(1) = 1$, we start establishing upper bounds at $g(2)$.

Lemma 4.1. *Duplicator can win 2-round MS games whenever both linear orders are of size 2 or greater, and hence $g(2) \leq 2$.*

Proof. In the Introduction we considered a 2-round MS game on two linear orders of sizes $|B| = 3, |L| = 2$. Figure 1 is given again here for ease of reference. The Introduction covered



Figure 2: The case $|B| = 3, |L| = 2$.

the case where Spoiler selects the middle B element, $B(2)$, in the first round. The case where Spoiler picks an end element from either the L or B boards in the first round is easier – Duplicator just picks the corresponding end element from the other linear order and she does not even need to make a second copy of the board to win. For example, in response to $B(1)$, Duplicator will play $L(1)$, or in response to $L(2)$, Duplicator will play $B(3)$, in either case leading to simple wins.

In the example given in the Introduction, where Spoiler played $B(2)$, once Duplicator makes copies and plays different moves in each copy, we render the game after round 1, as in Figure 3.

To complete the analysis, we must show that Duplicator wins whenever $2 \leq |L| < |B|$. Let us begin with the case $|L| = 2$ and $|B| > 3$. If Spoiler picks an end element from B or from L , then Duplicator picks the corresponding end element from the opposite side and wins. On the other hand, if Spoiler picks a non-end element on B , Duplicator picks $L(1)$ on top and $L(2)$ on bottom, winning just like in the introduction. We are left to consider



Figure 3: In response to Spoiler playing $B(2)$, Duplicator makes a second copy of the L board and place $L(1)$ on one board and $L(2)$ on the other board. (The diagrams omit parentheses.)

the case when $3 \leq |L| < |B|$. Playing end moves from either B or L have the same effect as before. While if Spoiler picks a non-end element from either B (or L), Duplicator wins by playing *any* non-end element from, respectively, L (or B), guaranteeing a 2-round win. \square

Lemma 4.2. *Duplicator can win 3-round MS games on linear orders whenever both linear orders are of size 4 or greater, and hence $g(3) \leq 4$.*

Proof. Let us start with the base case $|B| = 5, |L| = 4$. See Figure 4. Duplicator-winning



Figure 4: The case $|B| = 5, |L| = 4$.

outcomes associated with all Spoiler 1st round plays, *other than* $B(3)$, are easy to analyze and described in the table below. Further explanation of the notation used in the table is given in the text that follows it.

| Round 1 | | Round 2 | | Round 3 | |
|---------|--------|-----------------------------------|--------|-------------------------------------|---|
| S | D | S | D | | |
| $B(1)$ | $L(1)$ | D wins by reduction to $f(2)$ | | | |
| $B(2)$ | $L(2)$ | $B(1)$ | $L(1)$ | D wins | |
| | | $B(2)$ | $L(2)$ | D wins | |
| | | $B(3)$ | $L(3)$ | D wins | |
| | | $B(4)$ | $L(3)$ | D wins on this board or board below | |
| | | | $L(4)$ | | |
| | | | $B(5)$ | $L(4)$ | D wins |
| | | | $L(1)$ | $B(1)$ | D wins by transposition to Rd2. $B(1),L(1)$ |
| | | | $L(2)$ | $B(2)$ | D wins by transposition to Rd2. $B(2),L(2)$ |
| | | | $L(3)$ | $B(3)$ | D wins by transposition to Rd2. $B(3),L(3)$ |
| | | | $L(4)$ | $B(5)$ | D wins by transposition to Rd2. $B(5),L(4)$ |
| $B(3)$ | $L(1)$ | Described in the text | | | |
| | | $L(2)$ | " | " | |
| | | $L(3)$ | " | " | |
| | | $L(4)$ | " | " | |
| $B(4)$ | $L(3)$ | Symmetrical to Rd1. $B(2),L(2)$ | | | |
| $B(5)$ | $L(4)$ | Symmetrical to Rd1. $B(1),L(1)$ | | | |
| $L(1)$ | $B(1)$ | Transposition to Rd1. $B(1),L(1)$ | | | |
| $L(2)$ | $B(2)$ | Transposition to Rd1. $B(2),L(2)$ | | | |
| $L(3)$ | $B(4)$ | Transposition to Rd1. $B(4),L(3)$ | | | |
| $L(4)$ | $B(5)$ | Transposition to Rd1. $B(5),L(4)$ | | | |

A few notes on the table: Moves given in the respective **S** columns are Spoiler plays while those given in the **D** columns are Duplicator plays. When we say that a game is winnable for Duplicator “by reduction to $f(k)$,” for some k , we mean that the game from this point on can be won by Duplicator simply as a 1-board standard Ehrenfeucht-Fraïssé game. Recall the definition of f from Theorem 1.4. In these cases, typically an end element has been played, and from an Ehrenfeucht-Fraïssé point of view, it is of no benefit to Spoiler to subsequently play on top of the 1st element, so we may remove these elements from both L and B and consider the game, in subsequent rounds, to be played solely on the remaining elements. The game reduces to $f(k)$ if there are k rounds yet to be played and both sets of remaining elements are at least of size $f(k)$.

When we say that a given sequence of moves X, Y is “symmetrical” to another sequence of moves X', Y' , we mean that you can arrive at one sequence by taking the mirror image of the other sequence with respect to the center of the boards, where the other sequence has already been analyzed.

We say that a sequence of moves is a “transposition” of another sequence of moves if the end board positions are the same but the sequence of moves leading to that board position is different.

Note that if Duplicator is able to win in a single additional move by focusing just on a single board on the B and L sides without any special strategy, we do not describe every possible move and response sequence (there are just too many), but rather just indicate that “D wins.”

In response to the one tricky Spoiler 1st round move of $B(3)$, Duplicator creates additional copies of the L board and makes every possible move, as depicted in Figure 5. Let’s now analyze the possible Spoiler 2nd round responses.

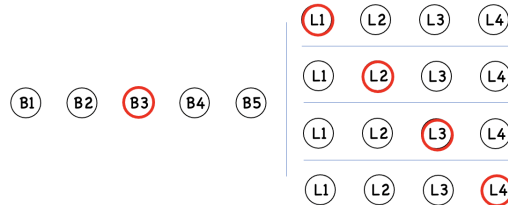


Figure 5: After Spoiler plays $B(3)$ from the B side, Duplicator plays $L(1)$, $L(2)$, $L(3)$ and $L(4)$ on different boards on the L side.

Case where Spoiler makes 2nd round move on B :

| Round 2 | |
|---------|--|
| Spoiler | Duplicator |
| $B(1)$ | $L(1)$ on 3rd L board insuring a win on that board |
| $B(2)$ | $L(2)$ on 3rd L board insuring a win on that board |
| $B(3)$ | $L(3)$ on 3rd L board insuring a win on that board |
| $B(4)$ | $L(3)$ on 2nd L board insuring a win on that board |
| $B(5)$ | $L(4)$ on 2nd L board insuring a win on that board |

Case where Spoiler makes 2nd round moves on L :

First consider the possibility of Spoiler playing atop an existing move. If he plays atop $L(2)$ on board 2, or $L(3)$ on board 3, then Duplicator will play atop $B(3)$, ensuring victory on the associated boards in another move. Analogously, if Spoiler plays on top of *both* $L(1)$ on the top board and $L(4)$ on the bottom board, then Duplicator can play atop $B(3)$ guaranteeing a win on one of these L boards or the other. Thus, we may assume Spoiler plays atop *at most* one of the existing moves, with that one move being atop either $L(1)$ or $L(4)$. Of the at least three boards in which Spoiler does *not* play on top of existing moves, the moves are either to the right of the existing moves, or to the left of the existing moves, and hence, there must be at least two played to one side or the other of the existing moves. Without loss of generality, assume that at least two of these moves are to the right of the existing moves.

Suppose that one of the moves to the right is on the 2nd board. An $L(3)$ move would be met by a $B(4)$ response by Duplicator, while an $L(4)$ move would be met by a $B(5)$ response, in either case leading to a single board win for Duplicator. Thus, we can assume that the two 2nd round moves to the right of the existing moves are on the 1st and 3rd boards – the only two boards we need to consider to conclude this bit of the analysis. The element $L(4)$ must be selected from board 3. If $L(2)$ is selected from board 1 then Duplicator wins by selecting $B(4)$: any 3rd round Spoiler move on B is parried on one of the two L boards, while if Spoiler plays his 3rd round from L , Duplicator can maintain an isomorphism with any move played on either the 1st or 3rd board. We are left to consider just the possibilities that Spoiler plays $L(3)$ or $L(4)$ for his 2nd round move on board 1. Suppose he plays $L(3)$. Duplicator can then win by playing $B(5)$: if Spoiler plays his 3rd move from B then $B(4)$ is met with $L(2)$ on the 1st board and any other B move is easily parried on the 3rd L board. On the other hand, if Spoiler plays his 3rd move from L then whatever he does on the 3rd L board can be matched with an isomorphism-preserving move on B . In the final case where Spoiler plays $L(4)$ for his 2nd round move on the 1st L board, Duplicator responds with $B(5)$, guaranteeing an isomorphism. It follows that Duplicator can always win the $|B| = 5, |L| = 4$, 3-round game.

To complete the argument we must show that Duplicator can also win when one or both of $|B| > 5$ and $|L| > 4$. Let us start by considering the case of $|B| > 5, |L| = 4$. The analysis for any initial Spoiler move where he plays either an end move or a next-to-end move is precisely as earlier. Consider all other moves on B to be “middle” moves. Any such middle move is parried just like in Figure 5, by playing $L(1)$, $L(2)$, $L(3)$, and $L(4)$ on the different boards on the L side. The only new wrinkle in this analysis occurs in the case where Spoiler first plays a middle move on B . We illustrate for $|B| = 6$ and where Spoiler’s first play is $B(3)$. He can now play $B(5)$ and Duplicator must take new precautions because she cannot win by playing on a single board. She can, however, play $L(3)$ on the first L board, covering further play on the right of $B(3)$ by Spoiler, while also playing, say, $L(3)$ on board 2, to cover potential play by Spoiler on the left of $B(3)$. The rest of the analysis is precisely as in the smaller $|B|$ case.

Finally, consider the case $|B| > |L| > 4$. If Spoiler plays an end element or next-to-end element, Duplicator follows suit, playing, respectively, an end element or next-to-end element from the same side on the opposite board, leading to an easy victory: if Spoiler plays on both sides of the 1st round move in subsequent rounds he will clearly lose, while playing just on one side reduces to a losing 2 round game for him. Similarly, if Spoiler plays a “middle” element, Duplicator can respond playing any middle element from the opposite side. Again,

if Spoiler plays on both sides of the 1st round move in subsequent rounds he will clearly lose, but playing just on one side reduces to a losing 2 round game for him.

That completes the argument that Duplicator can win an MS game of 3 rounds whenever the size of all boards is at least 4 and hence establishes the lemma. \square

At this point it is natural to suspect that one can build up upper bounds recursively in a relatively simple manner. However, such an approach runs into unexpected difficulties, as we describe in the next section.

5. INTERLUDE: WHY NAIVE RECURSION CANNOT BE USED TO BUILD UP DUPLICATOR WINNING STRATEGIES IN MS GAMES

It is worth pausing to understand why a simple idea to use recursion to build up Duplicator-winning strategies, and hence upper bounds on $g(r)$, fails. Via Lemma 4.2, we have established the fact that $g(3) \leq 4$ by showing that Duplicator wins 3-round MS games if the sizes of both linear orders are 4 or larger. It is tempting to try to use this fact to produce a Duplicator strategy for winning 4-round games using recursion. To understand the problematic logic, it suffices to consider a 4-round game on boards of sizes 9 and 10. The erroneous argument runs as follows. Suppose Spoiler plays $L(5)$ on his 1st move. Duplicator can then simply reply with the single move $B(5)$ (so the erroneous reasoning goes), as in Figure 6. Since there are five unplayed elements to the right of $B(5)$ and four unplayed



Figure 6: A simple attempt to arrive at a Duplicator-winning strategy for the 10 versus 9 game. The first round moves are given in red.

elements to the right of $L(5)$, Duplicator should now just be able to mimic Spoiler's moves at, or to the left of, $B(5)/L(5)$ and otherwise play moves to the right of $B(5)/L(5)$ as if it were a 3-round, 5 versus 4 game, which we know is winnable by Duplicator. In fact, as we will learn in Lemma 6.3, the 10 vs. 9 game is winnable by *Spoiler*. Hence this strategy does not work.

The reason the strategy doesn't work is that there is interaction between play on the two sides and there are moves to the left of the 5 vs. 4 sub-game that are more powerful for Spoiler (in the sense of breaking more to-that-point maintained partial isomorphisms) than any moves available in the 5 vs. 4 game. This additional power is achieved by Spoiler playing atop an already played move that is not part of the 5 vs. 4 sub-game at a critical juncture.

Indeed, for his 2nd round move, Spoiler will select $B(8)$, and as we know from the analysis of the 5 vs. 4 games, this will require Duplicator to make copies of the L board and play each of the possible moves to the right of $L(5)$, as depicted in Figure 7. At this point Spoiler will play on top of $L(5)$ on the 1st L board (a move that *wasn't available* in the 5 vs. 4 game), on $L(6)$ on the 2nd L board, on $L(9)$ on the 3rd L board, and on top of $L(9)$ on the 4th L board. See Figure 8. Spoiler is going to play his 4th and final round moves from B , but first let's see what happens in response to the various Duplicator 3rd round moves (we assume the oblivious strategy). The only move that would keep an isomorphism with the 1st L board is a play atop $B(5)$ – breaking isomorphisms with any L board but the top one.

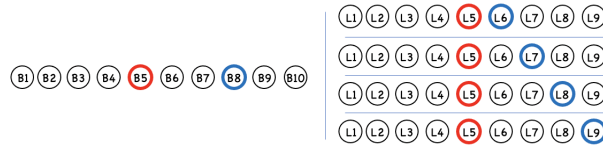


Figure 7: The 2nd round plays of the simple Duplicator-winning strategy for the 10 versus 9 game. The first round moves are given in red, 2nd round moves in blue.

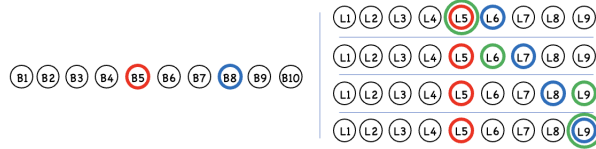


Figure 8: 3rd round plays (in green) that foils the simple Duplicator-winning strategy for the 10 versus 9 game. The first round moves are given in red, 2nd round moves in blue.

But then Spoiler will play $B(6)$ on his 4th move on the same board, and Duplicator will then not be able to keep the isomorphism going with the top L board. On the other hand, to maintain an isomorphism with the 2nd L board, Duplicator must play either $B(6)$ or $B(7)$, and in so doing, break an isomorphism with any other L board. However, Spoiler will respond with $B(7)$ if $B(6)$ was played, or $B(6)$ if $B(7)$ was played, and Duplicator will have no retort. To maintain an isomorphism with the 3rd L board, Duplicator must play $B(9)$ or $B(10)$, again breaking the isomorphism with all other L boards, and Spoiler will respond by playing $B(10)$ if $B(9)$ was played and vice versa. Finally, to keep an isomorphism going with the bottom L board, Duplicator must play on top of $B(8)$, again breaking the isomorphisms with other L boards. But then Spoiler plays $B(9)$ and Duplicator cannot respond.

Thus, trying to replicate the 5 vs. 4 strategy to the right of $B(5)/L(5)$ and mimicking play on top of or to the left of $B(5)/L(5)$ does not work for Duplicator. When Spoiler played on top of $L(5)$ for one of his 3rd round moves, he was utilizing a move that was *not* available to him in the 5 vs. 4 game. He *did* have the option to play on top of $L(6)$ (which was labeled $L(1)$ in the 5 vs. 4 game) – but doing so would have duplicated the isomorphism type of the bottom board. Playing on top of $L(5)$ has stronger effect since if Duplicator is to maintain isomorphisms with both the top and bottom boards, she cannot do it on a single board on the left, she must make a copy and maintain an isomorphism with the top-right board in one copy, and maintain an isomorphism with the bottom-right board in the other copy.

Before moving on, it is worth noting that it was not strictly necessary for Spoiler to play on top of $L(5)$ on the 1st L board on his 3rd move – any move on $L(1)–L(4)$ would have worked just as well. See Figure 9. For Duplicator to maintain an isomorphism with the top L board she will have to move on $B(1)–B(4)$, again breaking the isomorphisms with any other L boards, and Spoiler can then pick $B(6)$ on his 4th move, breaking any hope for an isomorphism with the 1st board.

We remark that every winning strategy for Spoiler in the 10 vs. 9 game requires at least one play-on-top move. We show this at the end of Appendix B, as summarized in Observation B.1.

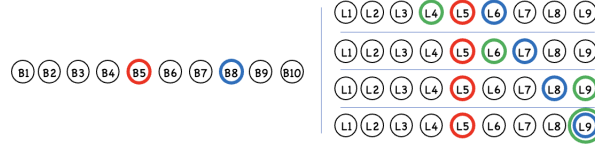


Figure 9: A second example of 3rd round plays (in green) that foils the simple Duplicator-winning strategy for the 10 versus 9 game. The first round moves are given in red, 2nd round moves in blue.

6. LOWER BOUNDS ON $g(r)$

The upshot of the prior section is that establishing tight bounds for $g(r)$ requires careful attention. Having established $g(1) = 1$ (Lemma 3.4), we begin our consideration of lower bounds with $g(2)$.

Lemma 6.1. $g(2) \geq 2$

Proof. The sentence $\Phi_2 = \exists x \exists y (x < y)$ distinguishes linear orders of size 2 and above, from the linear order of size 1. \square

Lemma 6.2. $g(3) \geq 4$

Proof. The following sentence, with 3 quantifiers, distinguishes linear orders of size at least 4 from those of size at most 3:

$$\Phi_3 = \forall x \exists y \exists z (x < y < z \vee y < z < x) \quad (6.1)$$

\square

Lemma 6.3. $g(4) \geq 10$.

Proof. The following sentence with 4 quantifiers distinguishes linear orders of size at least 10 from those of size at most 9.

$$\Phi_4 = \forall x \exists y \forall z \exists w ($$

$$x < z < y \rightarrow (w \neq z \wedge x < w < y) \wedge \quad (6.2)$$

$$x < y < z \rightarrow (w \neq z \wedge x < y < w) \wedge \quad (6.3)$$

$$y < z < x \rightarrow (w \neq z \wedge y < w < x) \wedge \quad (6.4)$$

$$z < y < x \rightarrow (w \neq z \wedge w < y < x) \wedge \quad (6.5)$$

$$z = x \rightarrow (x < w < y \vee y < w < x) \wedge \quad (6.6)$$

$$z = y \rightarrow (x < y < w \vee w < y < x)). \quad (6.7)$$

This sentence captures the fact that “for every x there is a y with two or more elements on each side of y , both of which are on the same side of x as y ”, a fact that is true for linear orders of size 10 or greater, but not for linear orders of size less than 10. For example, in a linear order of size 9, the middle element will not have an element to either side of it having these properties. More specifically, the first four implications (6.2)–(6.5) say that for every x there is a y such that for any $z \neq x, y$, with z on the same side of x as y , there is an additional element besides z on the same side of y and also on the same side of x as y . The last two implications (6.6)–(6.7) are critical and insure that there are elements (i) between x and y and (ii) less than y if $y < x$, and greater than y if $y > x$. \square

For an alternate, game-based proof of Lemma 6.3 see Appendix B.

With initial values $g(1) = 1, g(2) \geq 2$ (Lemma 6.1), $g(3) \geq 4$ (Lemma 6.2) and $g(4) \geq 10$ (Lemma 6.3), we establish all remaining lower bounds via a game argument:

Theorem 6.4. For $r > 4$,

$$g(r) \geq \begin{cases} 2g(r-1) & \text{if } r \text{ is even,} \\ 2g(r-1) + 1 & \text{if } r \text{ is odd.} \end{cases} \tag{6.8}$$

Proof. Let us establish (6.8) first for odd r . Here we need to provide a Spoiler-winning strategy for a game with arbitrary linear orders of sizes at least $2g(r-1) + 1$ on one side and arbitrary linear orders of sizes at most $2g(r-1)$ on the other side. For simplicity, we will show the strategy for the game in which we have a single linear order of size at least $2g(r-1) + 1$ on one side and a single linear order of size at most $2g(r-1)$ on the other side. The strategy for the case of multiple linear orders on each side is exactly the same, as we shall see. Analogous remarks will hold for the even r case.

We start with the special, though pivotal, case where we have linear orders of size $2g(r-1) + 1$ and $2g(r-1)$, as in Figure 10. We describe a winning strategy for Spoiler.

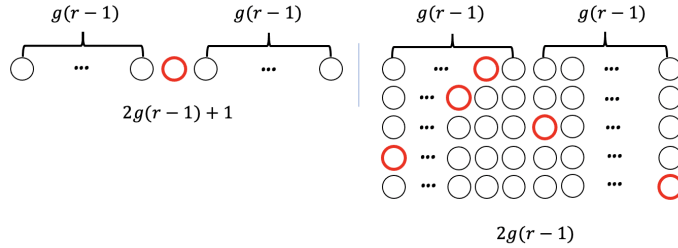


Figure 10: The odd r case: linear orders of sizes $2g(r-1) + 1$ and $2g(r-1)$. Spoiler plays first on B (the left hand side). First round moves are indicated in red. Duplicator responds by playing every possible move. Five exemplary boards with different moves played on each are shown.

Spoiler begins by playing the middle element on B , i.e. $B(g(r-1) + 1)$. Duplicator responds by playing every possible move leaving short sides that all have fewer than $g(r-1)$ elements. Spoiler next makes his 2nd round move on each board of L on the short sides, while playing on top of any end moves (moves with empty short sides), such as those shown on the bottom two boards in Figure 10. With the exception of the two end moves, which we will handle independently, Spoiler is going to play all subsequent moves entirely on the short sides of the boards of L and the corresponding sides of the boards of B – in other words if the short side of L is on the left, he will play on the left side of B , and vice versa – in accordance with the known Spoiler-winning strategy on boards of size $g(r-1)$ or greater vs. boards of size less than $g(r-1)$.

In order to see how he is able to do this, we need to use a strong inductive assumption, namely that in games of r rounds, when r is even, and there are linear orders of sizes $g(r)$ or greater on one side, and less than $g(r)$ on the other side, Spoiler can then always win by playing first on the L side. The game-based proof of Lemma 6.3 demonstrated such a strategy for the case $r = 4$, and we will have to keep this commitment when we cover

additional even r cases next. For now, though, we are assuming that r is odd, so $r - 1$ is even and we have such a strategy.

Duplicator will respond with the oblivious strategy, making many copies of B and playing all possible moves. From this point forward, the boards on both sides that have had their 2nd moves played to the right of the 1st moves conceptually constitute one game, and the boards on both sides that have had their 2nd moves played to the left of the 1st moves conceptually constitute a second game. The conceptual game played entirely on the left side of the boards can be won by Spoiler as well as the conceptual game played on the right, both times using the strong induction hypotheses. Note that since both games have the same larger size boards, Spoiler can win by picking the same side to play on (L or B) on all boards in each of the conceptual games, in each subsequent round, and hence the two conceptual games can be played round-by-round in tandem. In so doing, all partial isomorphisms are broken and so Spoiler wins the combined game.

However, we have not yet described how to take care of the case where Duplicator played end moves on L in the 1st round and Spoiler reciprocated by playing on top of these end moves. In order to maintain an isomorphism with either of these boards Duplicator will have to play on top of the 1st move on B – which will break any potential isomorphism with any other L boards. Now we use a key observation from the Spoiler-winning strategy in the 10+ vs. 9- game that established $g(4) \geq 10$ (see the proof of Lemma 6.3): in the next-to-last round Spoiler played on L , and in the last round he played on B . Note that the r -round Spoiler strategy recursively uses an $(r - 1)$ -round Spoiler strategy, eventually using the 4-round Spoiler strategy since the base case of this lemma is $r = 5$, which is defined in terms of $r = 4$. Thus, for boards that remain isomorphic long enough, Spoiler will play the last two rounds consecutively on L and then B .

In the next-to-last round, in which Spoiler plays on L , Spoiler will select any element to the right of $L(1)$ on the boards where $L(1)$ was played as a first move, and Spoiler will select any element to the left of $L(2g(r - 1))$ on the boards where $L(2g(r - 1))$ was played as a 1st move. As a result, to maintain partial isomorphisms with the boards in which, respectively, $L(1)$ and $L(2g(r - 1))$ were played, Duplicator will have to play so that the B boards that remain isomorphic with the board in which $L(1)$ was played are not isomorphic to the boards in which $L(2g(r - 1))$ was played, and vice versa. Thus, in the final round, Spoiler will play to the right of the middle element on the B boards that maintained a partial isomorphism with the boards where $L(2g(r - 1))$ was played first, thus killing all surviving partial isomorphisms, and will play to the left of the middle element on the B boards that maintained a partial isomorphism with boards where $L(1)$ was played first, killing all of its remaining partial isomorphisms.

For the more general case where $|B| > 2g(r - 1) + 1$, Spoiler simply picks any element having at least $g(r - 1)$ elements on each side of his 1st move and the play proceeds via the same induction. Analogously, if $|L| < 2g(r - 1)$ it still follows that Duplicator's 1st round moves leave a short side of size less than $g(r - 1)$, and hence the argument remains the same with respect to such smaller size L .

In the case where we start with multiple boards $\{B_i\}, \{L_j\}$ of sizes $|B_i| \geq 2g(r - 1) + 1$ and $|L_j| \leq 2g(r - 1)$, Spoiler starts by playing on each board on the $\{B_i\}$ side exactly as if there were just that one board there, and responds in turn, with the same alternating strategy as we have outlined in the single board case. Instead of making copies of just one board on her first round moves, Duplicator will make copies of all boards and make every possible move. Spoiler will then respond as before, playing on the short sides (or on top of

end elements if necessary), and the argument recurses as in the case where we started with single boards.

Next let us tackle the case where r is even. We begin as usual with the base case of boards of sizes $2g(r-1)$ and $2g(r-1)-1$. See Figure 11.

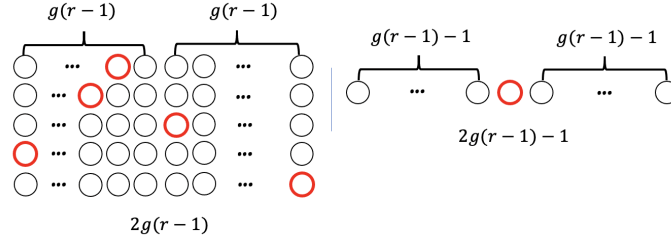


Figure 11: The even r case: linear orders of sizes $2g(r-1)$ and $2g(r-1)-1$. Spoiler plays first on L (the right hand side). First round moves are indicated in red. Duplicator responds by playing every possible move. Five exemplary boards with different moves played on each are shown.

As Spoiler, we must keep our inductive commitment to play first on L and so play the middle element, as indicated in the figure. Now any move by Duplicator on B leaves a long side of size at least $g(r-1)$ versus the same side (left or right of the 1st move) on L , of size $g(r-1)-1$. Thus we adapt the argument from the odd r case to this case, but where we play on the long side now rather than on the short side. The strong induction hypothesis we need this time is that smaller odd r cases can be won by playing the 1st round from B – but we know this to be the case for $r = 5$, the very first case covered by this theorem, and all subsequent cases of odd r by how we argued the odd r case. The case of boards of sizes greater than $2g(r-1)$ and less than $2g(r-1)-1$ is handled analogously, and the same remarks about multiple boards versus multiple boards to start that we made for the odd r case apply here as well. The theorem therefore follows. \square

For $r > 4$, the sentences associated with the strategies described in the proof of the above theorem say, respectively, for r even, that for every x there is a linear order of size at least $g(r-1)$ either to the left or right of x , and for r odd, that there is an element x with a linear order of size at least $g(r-1)$ both to the left and right of x .

We give a sketch of how this works for $r = 5$ and $r = 6$. To get started we rewrite the previous expression for Φ_4 , given by (6.2) – (6.7), replacing the variables x, y, z, w , respectively with x_2, x_3, x_4, x_5 :

$$\Phi_4 = \forall x_2 \exists x_3 \forall x_4 \exists x_5 ($$

$$x_2 < x_4 < x_3 \rightarrow (x_5 \neq x_4 \wedge x_2 < x_5 < x_3) \wedge \quad (6.9)$$

$$x_2 < x_3 < x_4 \rightarrow (x_5 \neq x_4 \wedge x_2 < x_3 < x_5) \wedge \quad (6.10)$$

$$x_3 < x_4 < x_2 \rightarrow (x_5 \neq x_4 \wedge x_3 < x_5 < x_2) \wedge \quad (6.11)$$

$$x_4 < x_3 < x_2 \rightarrow (x_5 \neq x_4 \wedge x_5 < x_3 < x_2) \wedge \quad (6.12)$$

$$x_4 = x_2 \rightarrow (x_2 < x_5 < x_3 \vee x_3 < x_5 < x_2) \wedge \quad (6.13)$$

$$x_4 = x_3 \rightarrow (x_2 < x_3 < x_5 \vee x_5 < x_3 < x_2)). \quad (6.14)$$

With this translation of variable names, the sentence says that “for every x_2 there is a x_3 with two or more elements on each side of x_3 , both of which are on the same side of x_2 as

x_3 ". More specifically, as noted before, the first four implications (6.9)–(6.12) say that for every x_2 there is an x_3 such that for any $x_4 \neq x_2, x_3$, with x_4 on the same side of x_2 as x_3 , there is an additional element besides x_4 on the same side of x_3 and also on the same side of x_2 as x_3 . The last two implications (6.13)–(6.14) insure that there are elements (i) between x_2 and x_3 and (ii) less than x_3 if $x_3 < x_2$, and greater than x_3 if $x_3 > x_2$.

To this we now add that there exists an element x_1 such that there is a linear order of size 10 both to its right and its left, in other words there exists an element x_1 such that the above sentence is true for elements less than x_1 and for elements greater than x_1 , specifically:

$$\begin{aligned} \Phi_5 = \exists x_1 \forall x_2 \exists x_3 \forall x_4 \exists x_5 \left(\right. \\ & x_1 < x_2 \rightarrow \left(x_1 < x_2 < x_4 < x_3 \rightarrow (x_5 \neq x_4 \wedge x_1 < x_2 < x_5 < x_3) \wedge \right. \\ & \quad x_1 < x_2 < x_3 < x_4 \rightarrow (x_5 \neq x_4 \wedge x_1 < x_2 < x_3 < x_5) \wedge \\ & \quad x_1 < x_3 < x_4 < x_2 \rightarrow (x_5 \neq x_4 \wedge x_1 < x_3 < x_5 < x_2) \wedge \\ & \quad x_1 < x_4 < x_3 < x_2 \rightarrow (x_5 \neq x_4 \wedge x_1 < x_5 < x_3 < x_2) \wedge \\ & \quad x_4 = x_2 \rightarrow (x_1 < x_2 < x_5 < x_3 \vee x_1 < x_3 < x_5 < x_2) \wedge \\ & \quad \left. x_4 = x_3 \rightarrow (x_1 < x_2 < x_3 < x_5 \vee x_1 < x_5 < x_3 < x_2) \right) \wedge \\ & x_2 < x_1 \rightarrow \left(x_2 < x_4 < x_3 < x_1 \rightarrow (x_5 \neq x_4 \wedge x_2 < x_5 < x_3 < x_1) \wedge \right. \\ & \quad x_2 < x_3 < x_4 < x_1 \rightarrow (x_5 \neq x_4 \wedge x_2 < x_3 < x_5 < x_1) \wedge \\ & \quad x_3 < x_4 < x_2 < x_1 \rightarrow (x_5 \neq x_4 \wedge x_3 < x_5 < x_2 < x_1) \wedge \\ & \quad x_4 < x_3 < x_2 < x_1 \rightarrow (x_5 \neq x_4 \wedge x_5 < x_3 < x_2 < x_1) \wedge \\ & \quad x_4 = x_2 \rightarrow (x_2 < x_5 < x_3 < x_1 \vee x_3 < x_5 < x_2 < x_1) \wedge \\ & \quad \left. x_4 = x_3 \rightarrow (x_2 < x_3 < x_5 < x_1 \vee x_5 < x_3 < x_2 < x_1) \right) \wedge \\ & \left. x_1 = x_2 \rightarrow (x_3 < x_1 \wedge x_1 < x_5) \right). \end{aligned}$$

Then Φ_5 distinguishes linear orders of size 21 and above from those of size less than 21.

Next, to form the expression for Φ_6 , which distinguishes linear orders of size 42 and above from those of size less than 42, we say that for every element x_0 there is either a linear order of size 21 to the left or right of x_0 . The construction is analogous:

$$\begin{aligned} \Phi_6 = \forall x_0 \exists x_1 \forall x_2 \exists x_3 \forall x_4 \exists x_5 \left(\right. \\ & x_0 < x_1 \rightarrow \left(x_0 < x_1 < x_2 \rightarrow \right. \\ & \quad \left(x_0 < x_1 < x_2 < x_4 < x_3 \rightarrow (x_5 \neq x_4 \wedge x_0 < x_1 < x_2 < x_5 < x_3) \wedge \right. \\ & \quad \quad x_0 < x_1 < x_2 < x_3 < x_4 \rightarrow (x_5 \neq x_4 \wedge x_0 < x_1 < x_2 < x_3 < x_5) \wedge \\ & \quad \quad x_0 < x_1 < x_3 < x_4 < x_2 \rightarrow (x_5 \neq x_4 \wedge x_0 < x_1 < x_3 < x_5 < x_2) \wedge \\ & \quad \quad x_0 < x_1 < x_4 < x_3 < x_2 \rightarrow (x_5 \neq x_4 \wedge x_0 < x_1 < x_5 < x_3 < x_2) \wedge \\ & \quad \quad x_4 = x_2 \rightarrow (x_0 < x_1 < x_2 < x_5 < x_3 \vee x_0 < x_1 < x_3 < x_5 < x_2) \wedge \\ & \quad \quad \left. x_4 = x_3 \rightarrow (x_0 < x_1 < x_2 < x_3 < x_5 \vee x_0 < x_1 < x_5 < x_3 < x_2) \right) \wedge \\ & \left. x_0 > x_1 \rightarrow \left(x_0 > x_1 > x_2 \rightarrow \right. \right. \\ & \quad \left(x_0 > x_1 > x_2 > x_4 > x_3 \rightarrow (x_5 \neq x_4 \wedge x_0 > x_1 > x_2 > x_5 > x_3) \wedge \right. \\ & \quad \quad x_0 > x_1 > x_2 > x_3 > x_4 \rightarrow (x_5 \neq x_4 \wedge x_0 > x_1 > x_2 > x_3 > x_5) \wedge \\ & \quad \quad x_0 > x_1 > x_3 > x_4 > x_2 \rightarrow (x_5 \neq x_4 \wedge x_0 > x_1 > x_3 > x_5 > x_2) \wedge \\ & \quad \quad x_0 > x_1 > x_4 > x_3 > x_2 \rightarrow (x_5 \neq x_4 \wedge x_0 > x_1 > x_5 > x_3 > x_2) \wedge \\ & \quad \quad x_4 = x_2 \rightarrow (x_0 > x_1 > x_2 > x_5 > x_3 \vee x_0 > x_1 > x_3 > x_5 > x_2) \wedge \\ & \quad \quad \left. x_4 = x_3 \rightarrow (x_0 > x_1 > x_2 > x_3 > x_5 \vee x_0 > x_1 > x_5 > x_3 > x_2) \right) \wedge \\ & \left. x_0 = x_1 \rightarrow (x_2 < x_0 \wedge x_0 < x_5) \right). \end{aligned}$$

$$\begin{aligned}
& x_0 < x_2 < x_1 \rightarrow \left(\begin{aligned}
& x_0 < x_2 < x_4 < x_3 < x_1 \rightarrow (x_5 \neq x_4 \wedge x_0 < x_2 < x_5 < x_3 < x_1) \wedge \\
& x_0 < x_2 < x_3 < x_4 < x_1 \rightarrow (x_5 \neq x_4 \wedge x_0 < x_2 < x_3 < x_5 < x_1) \wedge \\
& x_0 < x_3 < x_4 < x_2 < x_1 \rightarrow (x_5 \neq x_4 \wedge x_0 < x_3 < x_5 < x_2 < x_1) \wedge \\
& x_0 < x_4 < x_3 < x_2 < x_1 \rightarrow (x_5 \neq x_4 \wedge x_0 < x_5 < x_3 < x_2 < x_1) \wedge \\
& x_4 = x_2 \rightarrow (x_0 < x_2 < x_5 < x_3 < x_1 \vee x_0 < x_3 < x_5 < x_2 < x_1) \wedge \\
& x_4 = x_3 \rightarrow (x_0 < x_2 < x_3 < x_5 < x_1 \vee x_0 < x_5 < x_3 < x_2 < x_1) \end{aligned} \right) \wedge \\
& (x_0 < x_1 \wedge x_1 = x_2) \rightarrow (x_0 < x_3 < x_1 \wedge x_0 < x_1 < x_5) \\
& \wedge \\
& x_1 < x_0 \rightarrow \left(\begin{aligned}
& x_1 < x_2 < x_0 \rightarrow \left(\begin{aligned}
& x_1 < x_2 < x_4 < x_3 < x_0 \rightarrow (x_5 \neq x_4 \wedge x_1 < x_2 < x_5 < x_3 < x_0) \wedge \\
& x_1 < x_2 < x_3 < x_4 < x_0 \rightarrow (x_5 \neq x_4 \wedge x_1 < x_2 < x_3 < x_5 < x_0) \wedge \\
& x_1 < x_3 < x_4 < x_2 < x_0 \rightarrow (x_5 \neq x_4 \wedge x_1 < x_3 < x_5 < x_2 < x_0) \wedge \\
& x_1 < x_4 < x_3 < x_2 < x_0 \rightarrow (x_5 \neq x_4 \wedge x_1 < x_5 < x_3 < x_2 < x_0) \wedge \\
& x_4 = x_2 \rightarrow (x_1 < x_2 < x_5 < x_3 < x_0 \vee x_1 < x_3 < x_5 < x_2 < x_0) \wedge \\
& x_4 = x_3 \rightarrow (x_1 < x_2 < x_3 < x_5 < x_0 \vee x_1 < x_5 < x_3 < x_2 < x_0) \end{aligned} \right) \wedge \\
& x_2 < x_1 < x_0 \rightarrow \left(\begin{aligned}
& x_2 < x_4 < x_3 < x_1 < x_0 \rightarrow (x_5 \neq x_4 \wedge x_2 < x_5 < x_3 < x_1 < x_0) \wedge \\
& x_2 < x_3 < x_4 < x_1 < x_0 \rightarrow (x_5 \neq x_4 \wedge x_2 < x_3 < x_5 < x_1 < x_0) \wedge \\
& x_3 < x_4 < x_2 < x_1 < x_0 \rightarrow (x_5 \neq x_4 \wedge x_3 < x_5 < x_2 < x_1 < x_0) \wedge \\
& x_4 < x_3 < x_2 < x_1 < x_0 \rightarrow (x_5 \neq x_4 \wedge x_5 < x_3 < x_2 < x_1 < x_0) \wedge \\
& x_4 = x_2 \rightarrow x_0 < (x_2 < x_5 < x_3 < x_1 \vee x_3 < x_5 < x_2 < x_1 < x_0) \wedge \\
& x_4 = x_3 \rightarrow x_0 < (x_2 < x_3 < x_5 < x_1 \vee x_5 < x_3 < x_2 < x_1 < x_0) \end{aligned} \right) \wedge \\
& (x_1 = x_2 \wedge x_1 < x_0) \rightarrow (x_3 < x_1 < x_0 \wedge x_1 < x_5 < x_0) \\
& \wedge \\
& x_0 = x_1 \rightarrow x_3 \neq x_0 \end{aligned} \right).
\end{aligned}$$

It is worth noting that we could have begun this process at $g(4)$. Starting with the expression for Φ_3 , given by (6.1), establishing $g(3) \geq 4$ with prenex signature $\forall\exists\exists$ we could have “relativized” Φ_3 to form an expression with prenex signature $\exists\forall\exists\exists$ saying that there exists an element with a linear order of size at least 4 both to the left and right. This would have established that $g(4) \geq 9$. From there we could have obtained an expression with prenex signature $\forall\exists\forall\exists\exists$ establishing that $g(5) \geq 18$, and so on, at each juncture obtaining slightly worse bounds than we have already obtained. The magic is that a stronger lower bound for $g(4)$ can be expressed with the sentence given by (6.2) – (6.7) [alternatively, 6.9) – (6.14)].

7. GAMES WITH ATOMS: A DIFFERENT APPROACH TO OBTAIN UPPER BOUNDS ON $g(r)$

Let us define a new type of game. These games are similar to the MS games but with a twist. They will make it harder for Duplicator to win any particular game of r rounds. Since Duplicator-winning strategies for a game of r rounds provide upper bounds on $g(r)$, we will

obtain upper bounds that are potentially weaker, but will later match the lower bounds for $r \geq 4$, proving the upper bounds in the range $r \geq 4$ to be tight. The reason for considering these games is that they will allow us to recurse, getting around the issue we got stuck on in Section 5.

In our new game of r rounds, we again have sets \mathcal{A} and \mathcal{B} of structures. However, each board on each of the sides, in addition to containing a structure S , contains a collection of unrelated, but labeled elements, $\{a_1, \dots, a_s\}$, where s can be as large as Spoiler wants. We shall refer to the collection of unrelated labeled elements as *atoms*. By a slight abuse of notation, we will treat atoms as if they are constant symbols, so that they can appear in sentences. However, these symbols will not appear in the sentences guaranteed by Theorems 1.2 and Theorem 1.3. These atoms are both unrelated to elements of the structure and unrelated to each other. On their respective turns, Spoiler and Duplicator play, as in the MS games, on all boards on their chosen side for that turn, picking an element from that board, choosing either an element from the structure, or from the set of atoms. If atom a_k is selected by Spoiler on a given board, in a given round j , the only way Duplicator can maintain a partial isomorphism between that board and an opposite board is to select a_k in round j on the opposite board. When Duplicator makes copies of boards, the atoms are copied as well as the structures labeled with the moves made thus far.

In our figures, rather than showing all the atoms, and distinguishing those that are selected, we show only the atoms that have thus far been selected for a given board adjacent to the structures, with appropriate labeling. Let us refer to these new games as *MS games with atoms*.

Observation 7.1. MS games with atoms are equivalent to MS games with structures that are a union of the prior structures and the set of atoms. Hence, by Theorem 1.2, we have that Spoiler wins r -round MS games with atoms on two sets of structures iff there exists a sentence Φ with up to r quantifiers that distinguishes the two sets of structures. The sentence can contain constants and utilize an additional unary relation $A(x)$, which is true iff x is an atom.

Since a Spoiler winning strategy for a MS game is still a winning strategy for the analogous game with atoms (where atoms are never selected by Spoiler) we have:

Lemma 7.2. *If Spoiler can win an r -round MS game on a given pair of sets of structures then he can also win an r -round MS game with atoms on the same pair of sets of structures.*

Corollary 7.3. *If Duplicator can win an r -round MS game with atoms on a given pair of sets of structures, then she can also win an r -round MS game on the same pair of sets of structures.*

To understand the additional power provided to Spoiler by the atoms, let us revisit how Duplicator survives in one critical juncture of the 3-round, 5 versus 4 MS game from Lemma 4.2. Figure 12 shows the boards after two rounds in the critical sequence. In the 1st round, Spoiler played $B(3)$ and Duplicator played all possible moves on L . We consider the case where Spoiler then replies with 2nd round moves on L as indicated in blue in the figure, and in particular playing on top of $L(1)$ on the board labeled L_1 and on top of $L(4)$ on the board labeled L_4 . Duplicator then replies with all possible responses on B (again in blue). To the left of each B board we indicate which L boards the board has managed to keep an isomorphism with.

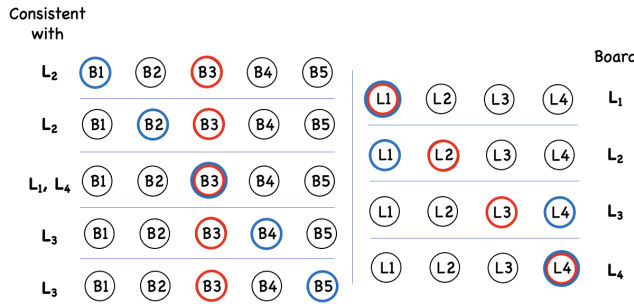


Figure 12: MS game play on \mathcal{B} versus \mathcal{L} . 1st round moves are in Red, 2nd round moves in blue.

The crux of the matter is that because the third B board is consistent with both L_1 and L_4 , Spoiler cannot break both isomorphisms in the one remaining move.

With atoms, however, investing a turn to play an atom can usefully separate these two L boards. Figure 13 illustrates Spoiler changing his second move on board L_4 from $L(4)$ to

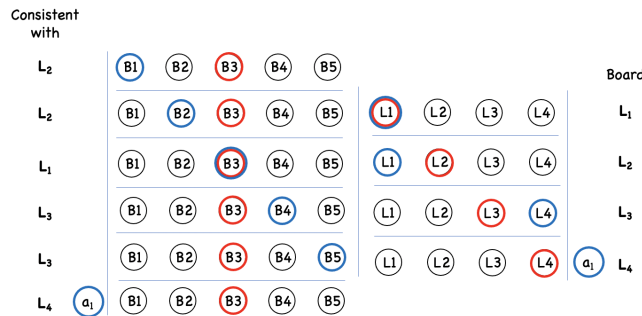


Figure 13: MS game play with atoms.

a_1 . Duplicator now has an additional potentially viable 2nd round move, playing the newly played atom, a_1 , as indicated in the additional board added on the left. However, note in the left hand column of the figure how no board is any longer simultaneously partially isomorphic with boards L_1 and L_4 . Spoiler can now break each of the isolated isomorphisms by playing on B , in order, from top to bottom, as follows: $B(2), B(1), B(2), B(5), B(4), B(4)$.

The following follows from the discussion above and Lemma 4.2.

Proposition 7.4. *While Duplicator can always win 3-round MS games on linear orders of sizes 5 vs. 4, Spoiler can always win 3-round MS games with atoms on linear orders of these sizes.*

We are now going to establish Duplicator-winning strategies for MS games with atoms, for various numbers of rounds. By Lemma 7.2, Duplicator has a harder time winning these games than standard MS games and so these strategies will initially provide weaker upper bounds than the upper bounds obtained by providing Duplicator-winning strategies in MS games.

Definition 7.5. Let $g'(r)$ denote the largest k such that Spoiler wins every r -round MS game with atoms on two sets \mathcal{B} and \mathcal{L} of linear orders where each $B \in \mathcal{B}$ is of size at least k and each $L \in \mathcal{L}$ is of size less than k .

To see that $g'(r)$ is well-defined, note that if Duplicator can win an r -round E-F game on two linear orders B, L then she can win an r -round MS game, as well as an r -round MS game with atoms, on sets \mathcal{B}, \mathcal{L} of structures with $B \in \mathcal{B}$ and $L \in \mathcal{L}$ – this is the case because she can focus on the one pair of structures, L and B , when playing the MS game, and further, because Spoiler gains no advantage from playing atoms when play is constrained to just a pair of structures. Hence $g'(r) \leq f(r)$. In our considerations of g' we will exclusively be focusing on game-based approaches. We will, further, just be concerned with establishing upper bounds on $g'(r)$, and in so doing will always be taking \mathcal{B} and \mathcal{L} to be singletons, $\mathcal{B} = \{B\}, \mathcal{L} = \{L\}$.

The following is an immediate consequence of Lemma 7.2.

Lemma 7.6. *For all values of r , we have $g(r) \leq g'(r)$.*

The following lemma describes why it is possible to recursively prove upper bounds on g' (and hence g) in MS games with atoms and get around the issue described in Section 5.

Lemma 7.7. Reduction Lemma: *Suppose we have an r -round MS game with atoms on linear orders of sizes K and K' . For some integer h , with $1 \leq h \leq \min(K, K')$, suppose there are boards on the L and B sides in which first moves of $L(h)$ and $B(h)$ have been played, or in which moves of $L(K - h + 1)$ and $B(K' - h + 1)$ have been played. Then Duplicator wins the K vs. K' game on those boards iff she wins the $(r - 1)$ -round MS game with atoms on linear orders of sizes $K - h$ and $K' - h$.*

Figure 14 illustrates the reduction of a 3-round, 6 versus 5 MS game with atoms, to a 2-round, 3 versus 2 MS game with atoms, per the conclusion.



Figure 14: An illustration of the reduction in Lemma 7.7 where a 3-round, 6 versus 5 MS game with atoms gets reduced to a 2-round, 3 versus 2 MS game with atoms.

Proof. Without loss of generality assume the 1st round moves are $L(h)$ and $B(h)$. Duplicator now makes a pact with Spoiler, saying that the only way she will maintain an isomorphism with moves on the left hand side of $L(h)/B(h)$ is by mirroring (in other words by responding to $L(i)$ with $B(i)$ when $1 \leq i \leq h$, and vice versa). She tells Spoiler that if ever such a move is *not* mirrored, Spoiler can count it as a break in the partial isomorphism between the two boards. By agreeing to these stricter isomorphism rules, Duplicator makes it harder for herself to maintain partial isomorphisms. However, the effect is that we can remove the elements $B(1), \dots, B(h)$ and $L(1), \dots, L(h)$ and set aside the first h atoms on each side, a_1, \dots, a_h , only to be played as follows: if Spoiler ever would have wanted to play $B(i)$ or $L(i)$, with $1 \leq i \leq h$, on a given board, he can instead play a_i with the same effect. This reduces the game to an $(r - 1)$ -round MS game with atoms on linear orders of size $K - h$ and $K' - h$ with h additional atoms. Since in the definition of MS games with atoms, Spoiler already has as many atoms as he wants, the additional h atoms have no effect on the game.

We are left with just playing an $(r - 1)$ -round MS game with atoms on linear orders of size $K - h$ and $K' - h$ and if Duplicator can win such a game, she can win the original game.

On the other hand, if Duplicator does *not* have a winning strategy in the $K - h$ vs. $K' - h$ MS game with atoms, then Spoiler has a winning strategy and can force play to be entirely on the side of the initial move with this many unplayed elements, and thus win. The lemma follows. \square

Lemma 7.8. *Laddering Up Lemma for MS Games and MS Games with Atoms:* *Suppose Duplicator can win MS games (respectively MS games with atoms) on singleton sets of sizes $K, K + 1$ for all $K \geq N$. Then Duplicator can also win MS games (respectively MS games with atoms) on singleton sets of sizes K, K' whenever $K \geq N$ and $K' \geq N$.*

Proof. Suppose both $K \geq N$ and $K' \geq N$. Let $\text{l.o.}(K)$ denote the linear order of size K . Then we have that $\text{l.o.}(K) \equiv_r \text{l.o.}(K + 1) \equiv_r \dots \equiv_r \text{l.o.}(K')$. By repeated application of Lemma 3.2 the lemma follows for MS games.

The same argument works for MS games with atoms by replacing each linear order with the union of the linear order and the corresponding atoms, whereby MS games with atoms reduce to MS games, per Observation 7.1. \square

Lemma 7.9. *Duplicator can win 2-round MS games with atoms on linear orders of sizes 2 or greater and hence $g'(2) \leq 2$.*

Proof. Immediate from Lemma 4.1 coupled with the observation that it never helps Spoiler to play an atom in the first or last round. \square

Lemma 7.10. *Duplicator can win 3-round MS games with atoms on linear orders of sizes 5 or greater and hence $g'(3) \leq 5$.*

Proof. Suppose we have linear orders of sizes $K, K + 1$ with $K \geq 5$. Any Spoiler 1st round move leaves a short side of no more than $\lfloor \frac{K}{2} \rfloor$ unplayed elements on that side. Duplicator can then reply with a move leaving an identical short side. Without loss of generality assume this common short side is on the left. Then to the right of the played moves, each board has at least $K - \lfloor \frac{K}{2} \rfloor - 1 \geq 2$ unplayed elements, and by virtue of the fact that $g'(2) \leq 2$ (Lemma 7.9), Lemma 7.7 guarantees that Duplicator has a winning strategy. The Laddering Up Lemma 7.8 then guarantees that Duplicator has a winning strategy for any boards of sizes $K \geq 5$ and $K' \geq 5$. The lemma follows. \square

Although this section is concerned with establishing upper bounds on g , we shall actually need to establish precise values for g' in order to get the upper bounds on g to go through. The discussion we gave to show that Spoiler can win a 5 vs. 4 3-round MS game with atoms (Proposition 7.4) can easily be extended to show that Spoiler can win such a 3-round game on linear orders of sizes 5 or greater versus 4 or smaller. Hence we have that $g'(3) \geq 5$, so that together with the prior lemma we have:

Lemma 7.11. $g'(3) = 5$.

Any Spoiler-winning strategy in an ordinary MS game of r rounds corresponds to a sentence that is valid for B but not for L with r quantifiers. If Spoiler's strategy starts on B then the corresponding sentence starts with \exists , while if Spoiler's strategy starts on L , the sentence starts with \forall . If we force Spoiler to play first on L , then we are giving Duplicator an advantage so that she may be able to win games that she would not be able to win otherwise.

Definition 7.12. Let $g'_\vee(r)$ denote the smallest k such that Duplicator can win MS games with atoms on a pair of linear orders, each of size k or greater, if Spoiler is constrained to play his first move on L .

Lemma 7.13. *If Spoiler is constrained to play his first move on L , Duplicator can win r -round MS games with atoms on linear orders of sizes $2g'(r-1)$ or greater. Hence, $g'_\vee(r) \leq 2g'(r-1)$.*

Proof. Consider an r -round, $2g'(r-1) + 1$ vs. $2g'(r-1)$ MS game with atoms. Refer to Figure 15. If Spoiler is constrained to play on the L side, his play will necessarily leave a

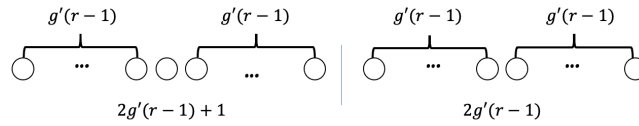


Figure 15: The r -round, $2g'(r-1) + 1$ vs. $2g'(r-1)$ MS game with atoms. Only one board on each side are shown.

short side of size at most $g'(r-1) - 1$, which can be matched by a move that leaves the same short side on B , and with long sides that are each of size at least $g'(r-1)$. Hence the game is winnable by Duplicator via the Reduction Lemma (Lemma 7.7). For r round games on boards of sizes $K + 1$ vs. K with $K > 2g'(r-1)$ again the short sides can be matched up, leaving long sides of sizes still at least $g'(r-1)$. The lemma follows by the Laddering Up Lemma (Lemma 7.8). \square

Theorem 7.14. *For $r \geq 2$,*

$$g'(r) = \begin{cases} 2g'(r-1) & \text{if } r \text{ is even,} \\ 2g'(r-1) + 1 & \text{if } r \text{ is odd.} \end{cases} \quad (7.1)$$

Moreover, Duplicator can win r -round MS games with atoms on linear orders of sizes $2g'(r-1)$ or greater if r is even, and on linear orders of sizes $2g'(r-1) + 1$ or greater if r is odd.

Proof. We establish the equality asserted in the theorem via first establishing that the \geq inequality holds, and then establishing that the \leq inequality holds. Since for all $r \geq 1$ we have $g'(r) \geq g(r)$ (lemma 7.6), the \geq portion of the theorem follows from the lemmas establishing that $g(1) = 1, g(2) \geq 2, g'(3) = 5$ and $g(4) \geq 10$, together with Theorem 6.4.

Now let us establish that the \leq inequality holds. It is trivial to verify that $g'(1) = 1$, while $g'(2) \leq 2$ is Lemma 7.9 and $g'(3) = 5$ is Lemma 7.11. Hence, we have already established the \leq part of the theorem for $r = 2$ and $r = 3$. For larger values of r we proceed inductively, assuming the truth of the theorem for values up to $r - 1$ and proving it for the value r . The inductive step for the case of odd r is easy, so let's dispose of that case first – it is essentially the same argument we gave to establish $g'(3) \leq 5$ in Lemma 7.10. Suppose we have linear orders of sizes $K, K + 1$ with $K \geq 2g'(r-1) + 1$. Any Spoiler 1st round move leaves a short side of no more than $\lfloor \frac{K}{2} \rfloor$ unplayed elements on that side. Duplicator can then reply with a move leaving an identical short side. Without loss of generality assume this common short side is on the left. Then, to the right of the played moves, each board has at least $K - \lfloor \frac{K}{2} \rfloor - 1$ unplayed elements. But, using the inductive assumption,

$K - \lfloor \frac{K}{2} \rfloor - 1 \geq K - \frac{K}{2} - 1 \geq (g'(r-1) + \frac{1}{2}) - 1$; hence $K - \lfloor \frac{K}{2} \rfloor - 1 \geq g'(r-1) - \frac{1}{2}$. Since both $K - \lfloor \frac{K}{2} \rfloor - 1$ and $g'(r-1)$ are integers, it follows that $K - \lfloor \frac{K}{2} \rfloor - 1 \geq g'(r-1)$. Thus, each board has at least $g'(r-1)$ unplayed elements on their long sides. The Reduction Lemma (Lemma 7.7) in conjunction with the induction hypothesis therefore guarantees that Duplicator has a winning strategy. The Laddering Up Lemma (Lemma 7.8) then guarantees that Duplicator has a winning strategy for any boards of sizes $K, K' \geq 2g'(r-1) + 1$.

Next consider the case of even r . Suppose first that we have linear orders of sizes $2g'(r-1)$ and $2g'(r-1) + 1$. Let us first dispose of any first move by Spoiler other than $B(g'(r-1) + 1)$, the middle element on the B side. Any move other than this one on the B side can be responded to with a move on L that matches the short side while leaving at least $g'(r-1)$ unplayed elements on both long sides, and therefore, again by the Reduction Lemma and induction, yielding a winning strategy for Duplicator. On the other hand, any Spoiler 1st round move on L is analogously met by matching the short side with a move on B , transposing to the just-prior analysis. Thus we may assume that Spoiler plays the element $B(g'(r-1) + 1)$. In response, Duplicator uses two boards and plays $L(g'(r-1))$ on one of the boards and $L(g'(r-1) + 1)$ on the other one. See Figure 16.

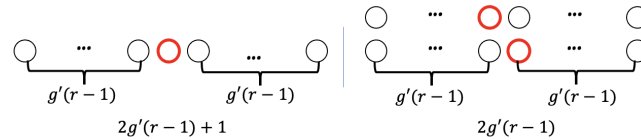


Figure 16: A $2g'(r-1) + 1$ vs. $2g'(r-1)$ game where Spoiler plays first on $B(g'(r-1) + 1)$ and Duplicator responds by playing $L(g'(r-1))$ on one board and $L(g'(r-1) + 1)$ on another.

Consider the possible Spoiler 2nd round responses. Suppose Spoiler plays on B . A play on one of the left-hand unplayed $g'(r-1)$ elements, i.e., on some $B(i)$ such that $1 \leq i \leq g'(r-1)$ will be met with a play of $L(i)$ on the 2nd L board. As we argued in the proof of the Reduction Lemma, we can now regard the remaining $g'(r-1) - 1$ unplayed elements to the left of the 1st round selections on the bottom L board and the B boards as additional atoms and just consider the game on the right side of these boards, which is now an $r - 2$ round game on boards of sizes $g'(r-1)$ and $g'(r-1) - 1$. Inductively it is easy to see that $g'(r-1) - 1 > g'(r-2)$ and so such a sequence leads inductively to a Duplicator win. A 2nd round Spoiler play on B , on one of the right hand set of unplayed $g'(r-1)$ elements, is handled with a symmetrical argument. Further, playing an atom on B is met by playing the same atom on both L boards (in fact playing on just one board and ignoring the other is sufficient), and again leads to an inductive win for Duplicator by virtue of the fact that $g'(r-1) - 1 > g'(r-2)$.

Thus we may assume that Spoiler makes his 2nd round moves on L . A play on the long side of either board is met with a symmetrical play by Duplicator on B , transposing to our prior analysis for when Spoiler played his 2nd move on B . Playing an atom on L is met with the same atom being played on B , again with a transposition. Thus we may suppose that Spoiler plays on the short side of both L boards, and, in particular, plays on the left on the top L board. Now the left hand (short side) of the top board is of size $g'(r-1) - 1$ and the left hand side of B is of size $g'(r-1)$. Further, we are assuming that r is even, and so

$r - 1$ is odd and hence by (7.1) we have

$$g'(r - 1) = 2g'(r - 2) + 1. \quad (7.2)$$

Since we've reduced the analysis to Spoiler next playing on the L side of this $(r - 1)$ -round, $g'(r - 1)$ vs. $g'(r - 1) - 1$ game, Lemma 7.13 applies and says that $g'_V(r - 1) \leq 2g'(r - 2)$. Taken together with (7.2), we have that $g'_V(r - 1) \leq g'(r - 1) - 1$.

Thus, we are left with boards of sizes $g'(r - 1)$ and $g'(r - 1) - 1$, both of which are at least of the size of $g'_V(r - 1)$. Hence, by the definition of g'_V (Definition 7.12), Duplicator has a winning strategy, over the remaining $r - 1$ rounds, playing just on the left hand sides of these two boards and hence, by the Reduction Lemma, has a winning strategy playing on the entire board. Thus Duplicator has a winning strategy in the original r -round game for boards of sizes $2g'(r - 1)$ and $2g'(r - 1) + 1$.

For $K, K + 1$ with $K > 2g'(r - 1)$ the argument is easier since Duplicator can just mimic the short side play of any 1st round Spoiler play and immediately apply the Reduction Lemma. As usual, the argument is completed by applying the Laddering Up Lemma. \square

We have therefore established the following upper bounds:

Corollary 7.15. *We have $g(2) \leq 2, g(3) \leq 4, g(4) \leq 10$, and for $r > 4$,*

$$g(r) \leq \begin{cases} 2g(r - 1) & \text{if } r \text{ is even,} \\ 2g(r - 1) + 1 & \text{if } r \text{ is odd.} \end{cases}$$

Moreover, Duplicator can win r -round MS games on linear orders of sizes that are at least as large as these upper bound (right hand side) values in all the inequalities.

Proof. The inequalities $g(2) \leq 2, g(3) \leq 4$ are Lemmas 4.1 and 4.2. The chain of inequalities $g(4) \leq g'(4) \leq 2g'(3) \leq 10$, follows from Lemma 7.6, Lemma 7.10 and Theorem 7.14. The inductively defined inequality for $r > 4$ follows by Lemma 7.6 and Theorem 7.14. Finally, the upper bounds associated with all these lemmas and Theorem 7.14 were established and stated by observing that Duplicator could win r round games when the two linear orders considered were at least as large as the upper bounds. The same therefore follows for this corollary. \square

8. TIGHT BOUNDS ON $g(r)$ AND FINAL THEOREMS

At last, we pull together the identical upper and lower bounds we have obtained for $g(r)$, to yield Theorem 1.6. Further, since the upper and lower bounds on $g(r)$ turned out to be tight, the very last paragraph at the end of Section 6 immediately implies the following:

Theorem 8.1. *For each $r \geq 1$ there is a sentence with r quantifiers distinguishing linear orders of size $g(r)$ or greater from linear orders of size less than $g(r)$. The prenex signatures of such sentences are as follows:*

$$\begin{aligned} r = 1 : & \quad \exists \\ r = 2 : & \quad \exists\exists \\ r = 3 : & \quad \forall\exists\exists \\ r \geq 4, r \text{ even} : & \quad \forall\exists \cdots \forall\exists \\ r \geq 4, r \text{ odd} : & \quad \exists\forall\exists \cdots \forall\exists. \end{aligned}$$

Here the \dots signifies a sequence of repeating quantifier alternations $\forall\exists$ of the length needed to give rise to r quantifiers in total.

Corollary 8.2. *Let \mathcal{A} be any collection of linear orders, each of size at least $g(r)$, and \mathcal{B} any collection of linear orders, each of size less than $g(r)$. The Spoiler can win an r -round MS game on \mathcal{A} and \mathcal{B} .*

Although we have a completely specified $g(r)$, we have not completely answered the question of the minimum number of quantifiers needed to distinguish one linear order, B , from another, L , in one special case. Specifically, if $g(r-1) \leq |L| < |B| < g(r)$ for some value of r , we know there is no sentence with $r-1$ quantifiers that distinguishes B from L , but nothing more. The following lemma closes this gap.

Lemma 8.3. *Given two linear orders B, L , with $|L| < |B| < g(r)$ for some $r > 0$ then there is a sentence with r quantifiers that distinguishes B from L .*

Proof. If there is a sentence with $r-1$ quantifiers that distinguishes B from L , there is obviously a sentence with r quantifiers that does the same. Hence, we only need to verify the lemma for the case $g(r-1) \leq |L| < |B| < g(r)$. This is the range of values of $|L|$ and $|B|$ for which we have not resolved how many quantifiers suffice to separate B from L .

We will demonstrate the conclusion of the lemma using a combination of explicit sentences and Spoiler winning strategies. It is not possible to have $|L| < |B| < g(1)$, so that case is handled. For $g(2)$ the only possibility is $|L| = 0, |B| = 1$, which is covered by $g(1) = 1$. Next, for $g(3)$, the cases covered by the lemma are those given in the table below.

| $ L $ | $ B $ | Why Spoiler Wins in 3 Rounds |
|-------|-------|--|
| 0 | 1 | $g(1) = 1$ |
| 0 | 2 | $g(1) = 1$ |
| 0 | 3 | $g(1) = 1$ |
| 1 | 2 | $g(2) = 2$ |
| 1 | 3 | $g(2) = 2$ |
| 2 | 3 | 3 rounds; Spoiler wins by playing 3 distinct elements in B |

Next consider $g(4)$, which we shall consider as a special case. Everything else will follow via an induction argument based on $g(r)$ for $r > 4$. Since $g(4) = 10$, we have $|B| \leq 9$. The following expression with 4 quantifiers distinguishes linear orders of size 9 and above from those of size 8 and below:

$$\begin{aligned} \Phi_{4,9} = \forall x \exists y \forall z \exists w (\\ & x < z < y \rightarrow (w \neq z \wedge x < w < y) \wedge \\ & x < y < z \rightarrow (w \neq z \wedge x < y < w) \wedge \\ & y < z < x \rightarrow (w \neq z \wedge y < w < x) \wedge \\ & z = x \rightarrow (x < w < y \vee y < w < x) \wedge \\ & z = y \rightarrow (x < y < w \vee w < y < x)). \end{aligned}$$

The sentence $\Phi_{4,9}$ is constructed from the sentence Φ_4 used in the proof of Lemma 6.3 by removing Condition (6.5) of Φ_4 . In words, $\Phi_{4,9}$ states that “for every x either there are 5 elements after it or 4 elements before it”. In more detail, “for every x there is a y such that if $y > x$ then there are two or more elements on each side of y , both of which are greater

than x and if $y < x$ then there are two or more elements between y and x and one element smaller than y ".

It is similarly easy to construct $\Phi_{4,8}$, $\Phi_{4,7}$, $\Phi_{4,6}$ and $\Phi_{4,5}$ just by removing more conditions from Φ_4 . We present these sentences below for completeness.

$$\begin{aligned} \Phi_{4,8} = \forall x \exists y \forall z \exists w (& \\ & x < y < z \rightarrow (w \neq z \wedge x < y < w) \wedge \\ & y < z < x \rightarrow (w \neq z \wedge y < w < x) \wedge \\ & z = x \rightarrow (x < w < y \vee y < w < x) \wedge \\ & z = y \rightarrow (x < y < w \vee w < y < x)). \end{aligned}$$

$$\begin{aligned} \Phi_{4,7} = \forall x \exists y \forall z \exists w (& \\ & x < y < z \rightarrow (w \neq z \wedge x < y < w) \wedge \\ & z = x \rightarrow (x < w < y \vee y < w < x) \wedge \\ & z = y \rightarrow (x < y < w \vee w < y < x)). \end{aligned}$$

$$\begin{aligned} \Phi_{4,6} = \forall x \exists y \forall z \exists w (& \\ & z = x \rightarrow (x < w < y \vee y < w < x) \wedge \\ & z = y \rightarrow (x < y < w \vee w < y < x)). \end{aligned}$$

$$\begin{aligned} \Phi_{4,5} = \forall x \exists y \forall z \exists w (& \\ & z = x \rightarrow (x < w < y \vee y < w < x) \wedge \\ & z = y \rightarrow (x < y < w \vee y < x)). \end{aligned}$$

We have ‘‘filled in the gaps’’ for our base case of $r = 4$ and we will now proceed by induction as we did in Theorem 6.4. A critical point is that all of these sentences for $\Phi_{4,k}$ for $4 < k \leq 9$ end with a universal and then an existential quantifier, meaning that Spoiler plays the next-to-last round on L and the last round on B .

Suppose then that $g(r - 1) \leq |L| < |B| < g(r)$ for $r > 4$. Spoiler adopts the same strategy as that described in the proof of Theorem 6.4 for the $g(r)$ vs $g(r) - 1$ game. One very minor nuance is that in the even r case there is not always a central element for Spoiler to select for his 1st round move in L since $|L|$ could be even. Similarly in the odd r case there is not always a central element for Spoiler to select for his 1st round move in B since $|B|$ could be even. It suffices, however, for Spoiler to play as close to the center as possible. As an example, the sentence $\Phi_{4,9}$ corresponds to playing the left point among the two middle points in L in the 9 versus 8 game. All other details of the argument are unchanged, including the fact that the last two rounds are played on L and then B . \square

Finally, we are able to prove the precise game theoretic analog of Theorem 1.7.

Theorem 8.4. *Two linear orders, B and L , with $|L| < |B|$, can be distinguished by a sentence with r quantifiers iff $|L| < g(r)$.*

Proof. If direction: If $|L| < |B| < g(r)$ then B and L can be distinguished by Lemma 8.3. If $|L| < g(r) \leq |B|$ then B and L can be distinguished by the definition of g (Definition 1.5).

Only if direction: If $g(r) \leq |L| < |B|$ then B and L cannot be distinguished by Corollary 7.15. \square

9. CONCLUSIONS

We have studied multi-structural (MS) games, which generalize E-F games by being played over sets \mathcal{A} , \mathcal{B} of structures rather than over a pair A , B of individual structures. Whereas E-F games can capture exactly the quantifier rank needed to describe a property, we showed that MS games can capture exactly the number of quantifiers needed to describe a property. As a first application, we used them to determine the number of quantifiers needed to distinguish between linear orders of different sizes. This determination became quite nuanced as we uncovered several pitfalls of interest in their own right.

The quantifier count is a natural complexity measure but has received scant attention compared to the quantifier rank, number of distinct variable names, and measures of size and depth of the sentence body.

We expect complexity differences to be magnified in studying other structures beyond linear orders, such as higher-dimensional lattices, rooted trees, and other classes of graphs. As with the related ideas of Lotfallah [Lot04], MS games extend readily to second-order logic where they may bear on higher-order problems in descriptive complexity. For example, MS games when adapted to second order logic can easily simulate Ajtai-Fagin games.

ACKNOWLEDGMENTS

We are indebted to Phokion Kolaitis for making Observation 3.3. We thank Neil Immerman and Leonid Libkin for helpful discussions.

REFERENCES

- [AF90] Miklós Ajtai and Ronald Fagin. Reachability is harder for directed than for undirected finite graphs. *J. Symbolic Logic*, 55(1):113–150, 03 1990. doi:10.2307/2274958.
- [AI03] Micah Adler and Neil Immerman. An $n!$ lower bound on formula size. *ACM Trans. Comput. Logic*, 4(3):296–314, July 2003. doi:10.1145/772062.772064.
- [Daw21] Anuj Dawar. Private communication, January 2021.
- [DS21] Anuj Dawar and Abhisekh Sankaran. Extension Preservation in the Finite and Prefix Classes of First Order Logic. In Christel Baier and Jean Goubault-Larrecq, editors, *29th EACSL Annual Conference on Computer Science Logic (CSL 2021)*, volume 183 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 18:1–18:13, Dagstuhl, Germany, 2021. Schloss Dagstuhl–Leibniz-Zentrum für Informatik. doi:10.4230/LIPIcs.CSL.2021.18.
- [Ehr61] Andrzej Ehrenfeucht. An application of games to the completeness problem for formalized theories. *Fundamenta Mathematicae*, 49:129–141, 1961. doi:10.4064/fm-49-2-129-141.
- [Fag75] Ronald Fagin. Monadic generalized spectra. *Mathematical Logic Quarterly*, 21(1):89–96, 1975. doi:10.1002/malq.19750210112.
- [Fag93] Ronald Fagin. Finite-model theory - a personal perspective. *Theoretical Computer Science*, 116(1):3 – 31, 1993. doi:10.1016/0304-3975(93)90218-I.
- [FLRV21a] Ronald Fagin, Jonathan Lenchner, Kenneth W. Regan, and Nikhil Vyas. Multi-structural games and number of quantifiers. *Proceedings of Logic in Computer Science (LICS)*, pages 1–13, 2021. doi:10.1109/LICS52264.2021.9470756.
- [FLRV21b] Ronald Fagin, Jonathan Lenchner, Kenneth W. Regan, and Nikhil Vyas. Multi-structural games and number of quantifiers. *CoRR*, 2021. <https://arxiv.org/abs/2104.14709v1>. doi:10.48550/arXiv.2104.14709.

- [Fra54] Roland Fraïssé. Sur quelques classifications des systèmes de relations. *Université d'Alger, Publications Scientifiques, Série A*, 1:35–182, 1954. doi:10.2307/2963939.
- [FSV95] Ronald Fagin, Larry J. Stockmeyer, and Moshe Y. Vardi. On monadic NP vs. monadic co-NP. *Inf. Comput.*, 120(1):78–92, 1995. doi:10.1006/inco.1995.1100.
- [GS05] Martin Grohe and Nicole Schweikardt. The succinctness of first-order logic on linear orders. *Log. Methods Comput. Sci.*, 1(6):1–25, 2005. doi:10.2168/LMCS-1(1:6)2005.
- [HV15] Lauri Hella and Jouko Väänänen. The size of a formula as a measure of complexity. In Åsa Hirvonen, Juha Kontinen, Roman Kossak, and Andrés Villaveces, editors, *Logic Without Borders - Essays on Set Theory, Model Theory, Philosophical Logic and Philosophy of Mathematics*, volume 5 of *Ontos Mathematical Logic*, pages 193–214. De Gruyter, 2015. doi:10.1515/9781614516873.193.
- [HV19] Lauri T Hella and Miikka S Vilander. Formula size games for modal logic and μ -calculus. *Journal of Logic and Computation*, 29(8):1311–1344, 12 2019. arXiv:<https://academic.oup.com/logcom/article-pdf/29/8/1311/32008427/exz025.pdf>, doi:10.1093/logcom/exz025.
- [Imm79] Neil Immerman. Length of predicate calculus formulas as a new complexity measure. In *20th Annual Symposium on Foundations of Computer Science, San Juan, Puerto Rico, 29-31 October 1979*, pages 337–347. IEEE Computer Society, 1979. doi:10.1109/SFCS.1979.21.
- [Imm81] Neil Immerman. Number of quantifiers is better than number of tape cells. *J. Comput. Syst. Sci.*, 22(3):384–406, 1981. doi:10.1016/0022-0000(81)90039-8.
- [Imm99] Neil Immerman. *Descriptive Complexity*. Graduate Texts in Computer Science. Springer, 1999. doi:10.1007/978-1-4612-0539-5.
- [Imm21] Neil Immerman. private communications, December 2020 and April 2021.
- [Lib12] Leonid Libkin. *Elements of Finite Model Theory*. Texts in Theoretical Computer Science. Springer, 2012. URL: <https://www.springer.com/gp/book/9783540212027>.
- [Lot04] Wafik Boulos Lotfallah. An Ehrenfeucht-Fraïssé class game. *Mathematical Logic Quarterly*, 50(2):179–188, 2004. doi:10.1002/malq.200310088.
- [LY05] Wafik Boulos Lotfallah and Maha Ragab Youssef. On the freedom of Spoiler in Ehrenfeucht-Fraïssé games. Available at [http://math.guc.edu.eg/Wafik/On the freedom of Spoiler in the EF game.doc](http://math.guc.edu.eg/Wafik/On%20the%20freedom%20of%20Spoiler%20in%20the%20EF%20game.doc), 2005.
- [Ros82] Joseph G. Rosenstein. *Linear Orderings*, volume 98 of *Pure and Applied Mathematics*. Academic Press, New York and London, 1982.
- [Ros05] Eric Rosen. On the first-order prefix hierarchy. *Notre Dame J. Formal Logic*, 46(2):147–164, 04 2005. doi:10.1305/ndjfl/1117755146.

APPENDIX A. A PROOF OF THEOREM 1.4

Proof. Note for future use that $f(1) = 1$ and $f(r) = 2f(r-1) + 1$ for $r > 1$. The statement is clearly true for $r = 1$. Assume inductively that it is true for $r - 1$. Let us refer to the bigger linear order as B (for “big”) and the smaller linear order by L (for “little”).

We first show that if the size of L is less than $f(r)$, then Spoiler wins the r -round game. There are two possibilities, depending on whether the size of B is odd or even. Assume first that it is odd, say of size $2k + 1$. So the size of L is at most $2k$, and since the size of L is less than $f(r)$, it follows that $2k < f(r) = 2f(r-1) + 1$, so $k \leq f(r-1)$. In the first round, Spoiler selects the middle point of B , call it s_1 . There are then k points to the left and k to the right of s_1 in B . After Duplicator selects a point d_1 in L , there will be either fewer than k points to the left of d_1 in L or fewer than k points to the right of d_1 in L .

In the former case, Spoiler now makes all of his moves to the left in either structure (forcing Duplicator to do the same in the other structure), and in the latter case (when there less than k points to the right of d_1 in L), Spoiler makes all of his moves to the right in either structure (forcing Duplicator to do the same in the other structure). Spoiler now wins by the inductive assumption, since $k \leq f(r-1)$, and Spoiler has turned this into a game with less than k elements in the smaller linear order.

Assume now that the size of B is even, say of size $2k$. So the size of L is at most $2k - 1$, and since the size of L is less than $f(r)$, we have

$$2k - 1 < f(r) = 2f(r-1) + 1,$$

so $2k - 2 < 2f(r-1)$, and so $k - 1 < f(r-1)$.

In the first round, Spoiler selects the k th point of B , call it s_1 . There are then $k - 1$ points to the left and k to the right of s_1 in B . Duplicator now selects a point d_1 in L . If d_1 is within the first $k - 1$ points in L , then Spoiler makes all remaining moves to the left in both structures, since there are $k - 1$ points to the left of s_1 , fewer than $k - 1$ points to the left of d_1 , and $k - 1 < f(r-1)$, so Spoiler wins by inductive hypothesis. If d_1 is not within the first $k - 1$ points in L , then there are at most $k - 1$ points to the right of d_1 . But there are k points to the right of s_1 . Spoiler makes all remaining moves to the right in either structures, and wins by inductive hypothesis since $k - 1 < f(r-1)$.

We now show that if the size of L is at least $f(r)$, then Duplicator wins the r -round game. If Spoiler selects s_1 within the first $f(r-1) + 1$ points in B or L , then Duplicator selects d_1 in the other structure with $d_1 = s_1$. There are now at least $f(r-1)$ points to the right of the first move in L , and more points than that to the right of the first move in B . Since Duplicator can simply mimic Spoiler’s choices when Spoiler moves to the left of the first point chosen, Duplicator wins by the inductive hypothesis in considering moves on the right.

Now assume that Spoiler selects s_1 beyond the first $f(r-1) + 1$ points in B or L . There are two cases, depending on whether Spoiler moves in B or L . Assume first that Spoiler moves in B , and selects point s_1 , which is the k th point from the right of B . This splits into two subcases, depending on whether or not $k \leq f(r-1) + 1$. If $k \leq f(r-1) + 1$, then Duplicator selects d_1 in L that is k points from the right of L . There are then at least $f(r-1)$ points in L to the left of d_1 . Duplicator simply mimics the moves of Spoiler on points to the right of s_1 or d_1 , and uses her winning strategy for points to the left of s_1 and d_1 , where there is a winning strategy by inductive assumption, since there at least $f(r-1)$ to the left of d_1 and more than that to the left of s_1 .

We now consider subcase 2, where $k > f(r - 1) + 1$. Let I (respectively, I') be the closed interval in B (respectively, L) where the left endpoint is at position $f(r - 1) + 1$ from the left of B (respectively, L), and whose right endpoint is at position $f(r - 1) + 1$ from the right of B (respectively, L). By assumption, the point s_1 in B is inside of I . The interval I' contains the point that is $f(r - 1) + 1$ from the left of L , and so is nonempty. Assume that the point s_1 in B is m from the left side of I and n from the right side of I . Since L is smaller in size than B , it follows that I' is smaller in size than I . So there is a point d_1 in I' that is at most m from the left side of I' , and at most n from the right side of I' . Since there are at least $f(r - 1)$ points to the left of d_1 in L and at least $f(r - 1)$ points to the right of d_1 in L , and since the number of points to the left (respectively, to the right) of d_1 in L is at most the number of points to the left (respectively, to the right) of s_1 in B , Duplicator can win by making use of her winning strategy on moves to the left of s_1 or d_1 , and make use of her winning strategy on moves to the right of s_1 and d_1 .

Assume now that Spoiler selects point s_1 in L . By assumption, s_1 is beyond the first $f(r - 1) + 1$ points in L . Assume s_1 is the k th point from the right of L . Then Duplicator selects d_1 as the k th point from the right of B . On moves to the right of d_1 or s_1 , Duplicator simply mimics Spoiler's moves, and on moves to the left of d_1 or s_1 , Duplicator has a winning strategy by inductive assumption, since there are more than $f(r - 1)$ points to the left of s_1 in L , and more than that to the left of d_1 in B . So again, Duplicator wins. This concludes the proof of the theorem. \square

APPENDIX B. GAME BASED PROOF OF LEMMA 6.3

Proof of Lemma 6.3 via Games. We first consider the singleton 10 vs. 9 base case. We later consider the case where $|L| < 9$, followed by $|B| > 10$, and lastly the case of many linear orders $\mathcal{B} = \{B_i\}$, $\mathcal{L} = \{L_j\}$, of sizes $|B_i| \geq 10$, $|L_j| \leq 9$. In the 10 vs. 9 game, Spoiler will play first by playing the central element, $L(5)$, on the L side. Without loss of generality, Duplicator then plays all possible moves on the B side. See Figure 17 for an illustration.

Since there are 10 elements on each board on the B side, there will necessarily be more elements on one side or the other of any move on any particular board. Refer to the side that has more elements as the “long side” and the side with fewer elements as the “short side.” Spoiler will now make his 2nd round moves on B , playing as close as he can to the middle of the long side of every board. Since there are at least 5 elements on the long side of every board, there will necessarily be at least two unplayed elements to either side of Spoiler's 2nd round move on the long side of each board. See Figure 18.

Refer to these moves that have two unplayed elements to either side of them as “middle moves.” Duplicator then follows up, playing every possible move on the L side. The possible moves are depicted in blue in Figure 19. There is plainly no value to playing on top of the 1st round move here so we omit that option.

For his 3rd move, Spoiler will now play on L , replying to each possible play of Duplicator using the green moves in Figure 20. In order to keep an isomorphism with the elements in boards 1 or 5 of Figure 20, Duplicator will have to play on top of her 1st move on B . Note that this move breaks the isomorphism with all other L boards except board 1 if the B board had its long side on the right of the first move, and except board 5 if the B board had its long side on the left of the first move. Spoiler is going to play his 4th moves from the B boards. In case the second move on B was right of the first move (in other words the long

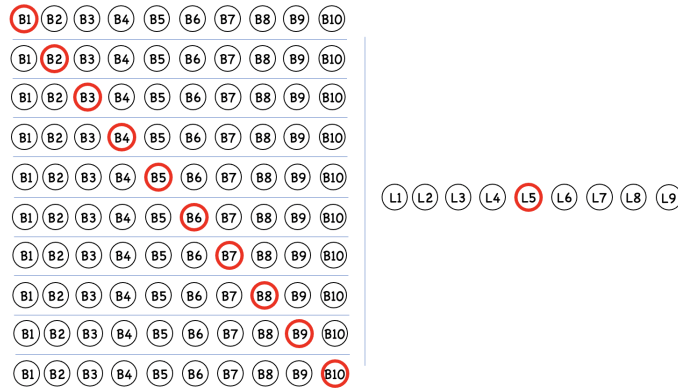


Figure 17: First Spoiler plays $L(5)$ on the L side. Duplicator responds by playing all possible moves on the B side. When discussing a given B board, given that there are 10 elements, there will either be more elements to the right or more elements to the left of the 1st move played. We refer to the side that has more elements as the “long side” and the side with fewer elements as the “short side.” Thus, in the figure, on the board in which Duplicator has played $B(5)$, the long side is to the right and the short side is to the left.

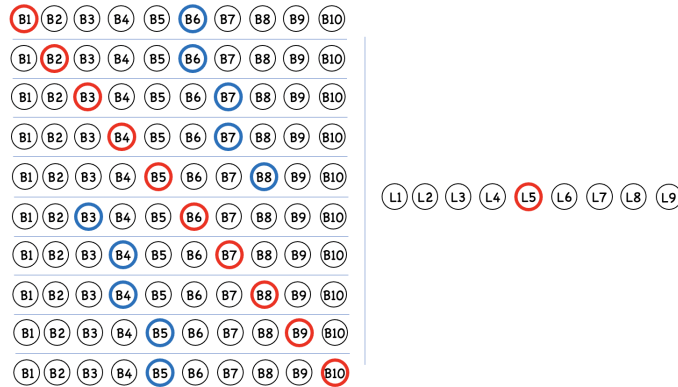


Figure 18: Spoiler’s 2nd moves on the B side (in blue), associated with each possible 1st round move on this side by Duplicator. The moves are all in the “middle” of the “long sides.” See the text for a description of these terms.

side was to the right), Spoiler will play to the right between the 1st and 2nd moves. Since there is no corresponding move on the 1st board, this move breaks the isomorphism with all L boards. Analogously, if the 2nd move on B was to the left of the 1st move, Spoiler will play to the left between the 1st and 2nd moves, again breaking the isomorphism with all boards. Since the arguments continue to be completely parallel, with plays on the top four boards of Figure 20 corresponding to cases where the long side of B was to the right, and play on boards 5 through 8 of Figure 20 corresponding to cases where the long side of B was to the left, we shall focus just on the top four boards. We have ruled out the case where Duplicator makes a 3rd round play that is on top of her 1st round play. To try to keep an isomorphism going with board 4 in the Figure, Duplicator must play on top of her 2nd move.

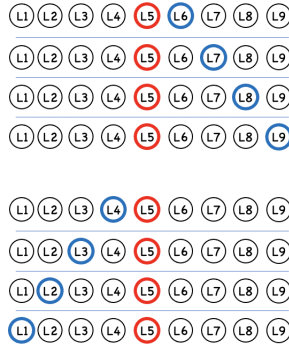


Figure 19: Possible 2nd moves by Duplicator (in blue) in response to Spoiler playing middle moves from the long side of all B boards. The play-on-top move is omitted, as discussed in the text.

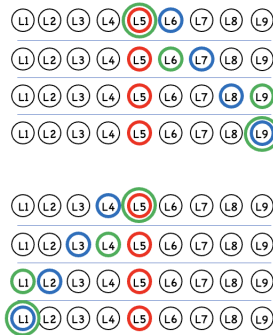


Figure 20: 3rd moves of Spoiler (in green) in response to each possible 2nd move of Duplicator (in blue).

But in this case Spoiler plays to the right of his 2nd move and there is no analogous move on board 4. Hence Duplicator cannot keep an isomorphism going with board 4. Next, to keep an isomorphism going with board 2 in Figure 20, Duplicator must play to the right between her 1st and 2nd moves. But then Spoiler plays a second time to the right between his 1st and 2nd moves, in the at least one additional unplayed element and Duplicator cannot reply in kind on the 2nd board. An attempt to maintain the isomorphism with board 3 in the Figure is seen to be fruitless with an analogous argument. Duplicator must play to the right of the 2nd move on B boards that have their long side to the right. Spoiler then picks a second unplayed element to the right of the 2nd move played on these boards and there is no analogous move on the 3rd board on the L side. It is thus impossible to maintain an isomorphism with any of the boards in Figure 20 and hence Spoiler wins this 10 versus 9 game.

Suppose now that $|L| < 9$. On his first move Spoiler will play a move that is as close to the center of L as possible. Duplicator will then just have fewer possible moves in the 2nd round when she plays on L because $|L|$ is smaller than in the prior analysis. Spoiler still exploits the fact that there are two unplayed elements both to the left and right of the 2nd round moves on the same side of the 1st round moves on B but not on L . If, for

a 2nd round move, Duplicator plays immediately to the right or immediately to the left of Spoiler's round 1 move, Spoiler will, on his 3rd round move, play on L on top of his 1st move (just as before). If Duplicator plays an end move, which is not simultaneously an immediate neighbor of Spoiler's 1st move, then Spoiler will play on top of that move (again, as before). On the other hand, if Duplicator plays a move that is two to the right (alternatively, two to the left) of the 1st round move, and the move is not also an end move, then Spoiler will play immediately to the right (alternatively, immediately to the left) of center. Analogously, a Duplicator move that is a 2nd-from-end element and not covered in any other case, will be responded to with an end move on the same side. In all cases the analysis is exactly as before and results in a Spoiler victory. Finally, if $|B| > 10$, the analysis is not particularly different than what we've seen; there is just a bit more "room" when picking Spoiler's 2nd round moves, which again must leave at least two unplayed elements both to the right and left. All other aspects of the analysis are unchanged. Finally, if we have a game with multiple linear orders, $\mathcal{B} = \{B_i\}, \mathcal{L} = \{L_j\}$ on the respective sides, with $|B_i| \geq 10$ and $|L_j| \leq 9$, then Spoiler first plays on each linear order in \mathcal{L} as described above, and then alternates – the fact that there happen to be multiple linear orders of different sizes has no impact on the strategy. The lemma is therefore established. \square

It is worth noting that Spoiler's ability to play on top of existing moves was essential to the above Spoiler-winning strategy. Suppose such a move were prohibited. Consider the situation after Duplicator's 2nd round moves – see Figure 19. Consider just the top 4 boards in this figure and suppose for the moment that our 3rd round moves were constrained to be on $L(6)$ – $L(9)$. By the analysis showing that the 5 vs. 4 game is Duplicator-winning (proof of Lemma 4.2), recall that playing two 3rd round moves with both moves either to the left of the 2nd round (blue) moves or to the right of the 2nd round moves would lead to a Duplicator victory. Thus, under the assumption that Spoiler does not play on top of an existing move, he would have to play moves on two boards in the $L(1)$ – $L(4)$ range. But then, for a 3rd round move, Duplicator could mimic any one of these Spoiler moves on the 4th B board (e.g., playing $B(i)$ if Spoiler played $L(i)$). It is then evident that any 4th round move on this 4th B board would be fruitless, while if Spoiler makes his 4th round moves on L , one of the moves on the two boards in which Spoiler previously played on $L(1)$ – $L(4)$ will allow for a Duplicator victory with respect to the same 4th B board.

It is worth calling out this last fact explicitly:

Observation B.1. If Spoiler were not able to play on top of existing moves, the M-S game on linear orders of sizes 10 vs. linear orders of size 9 would be winnable by Duplicator.