Journal of Computer and System Sciences **60**, 660–716 (2000)
doi:10.1006/jcss.1999.1691, available online at http://www.idealibrary.com on IDEAL®

The Closure of Monadic NP

Miklos Ajtai, Ronald Fagin, and Larry Stockmeyer

IBM Research Division, Almaden Research Center, 650 Harry Road, San Jose, California 95120-6099

E-mail: ajtai@almaden.ibm.com, fagin@almaden.ibm.com, stock@almaden.ibm.com

Received June 9, 1998

It is a well-known result of Fagin that the complexity class NP coincides with the class of problems expressible in existential second-order logic (Σ^1_1), which allows sentences consisting of a string of existential second-order quantifiers followed by a first-order formula. *Monadic NP* is the class of problems expressible in monadic Σ^1_1, i.e., Σ^1_1 with the restriction that the second-order quantifiers are all unary and hence range only over sets (as opposed to ranging over, say, binary relations). For example, the property of a graph being 3-colorable belongs to monadic NP, because 3-colorability can be expressed by saying that there exists three sets of vertices such that each vertex is in exactly one of the sets and no two vertices in the same set are connected by an edge. Unfortunately, monadic NP is not a robust class, in that it is not closed under first-order quantification. We define *closed monadic NP* to be the closure of monadic NP under first-order quantification and existential unary second-order quantification. Thus, closed monadic NP differs from monadic NP in that we allow the possibility of arbitrary interleavings of first-order quantifiers among the existential unary second-order quantifiers. We show that closed monadic NP is a natural, rich, and robust subclass of NP. As evidence for its richness, we show that not only is it a proper extension of monadic NP, but that it contains properties not in various other extensions of monadic NP. In particular, we show that closed monadic NP contains an undirected graph property not in the closure of monadic NP under first-order quantification and Boolean operations. Our lower-bound proofs require a number of new game-theoretic techniques. © 2000 Academic Press

1. INTRODUCTION

When faced with a very difficult mathematical question, such as the question of whether NP equals co-NP, one approach to answering the question is first to simplify the question, develop tools for solving the simpler version, and then try to extend the tools to move toward an answer to the original question.

Such an approach to the question of whether NP equals co-NP has spawned an area known as descriptive complexity, which is the complexity of describing

0022-0000/00 $35.00

problems in some logical formalism [Imm89]. This area began in the 1970s when Fagin [Fag74] showed that the complexity class NP coincides with the class of properties of finite structures expressible in existential second-order logic, otherwise known as Σ^1_1, which allows sentences consisting of a string of existential second-order quantifiers followed by a first-order formula.

One way of attacking the difficult question as to whether NP equals co-NP is to restrict the classes under consideration. Instead of considering Σ^1_1 (=NP) and its complement Π^1_1 (=co-NP) in their full generality, we could consider the monadic restriction of these classes, i.e., the restriction obtained by allowing second-order quantification only over sets (as opposed to quantification over, say, binary relations). We refer to the restricted classes as *monadic NP* and *monadic co-NP* [FSV95]. (It should be noted that, in spite of its restricted syntax, monadic NP does contain NP-complete problems, such as 3-colorability and satisfiability.) The hope is that the restriction to the monadic classes will yield more tractable questions and will serve as a training ground for attacking the problems in their full generality.

This line of attack was pursued by Fagin in [Fag75], where he separated monadic NP from monadic co-NP. Specifically, he showed that connectivity (of undirected graphs) is not in monadic NP, although it is easy to see that it is in monadic co-NP. Since then, monadic NP has been studied by a number of authors. For example, de Rougement [dR87], Fagin *et al.* [FSV95], and Schwentick [Sch95, Sch96] showed that connectivity is not in monadic NP, even in the presence of various built-in relations; Ajtai and Fagin [AF90] showed that directed (s, t)-connectivity is not in monadic NP, and their proof was simplified by Arora and Fagin [AF97]; Cosmadakis [Cos93] showed that a number of properties, including non-3-colorability, are not in monadic NP; Courcelle wrote a sequence of papers, including [Cou94], where he considered expressibility of monadic NP and of a variant where there is a different representation of graphs; Otto [Ott95] showed that monadic NP is strictly more expressive the more existential unary quantifiers are allowed; and Matz and Thomas [MT97b] and Schweikardt [Sch97] considered extensions of monadic NP obtained by alternating the unary second-order quantifiers.

One criticism of monadic NP is that it is not sufficiently robust. In particular, it is not closed under the simple operation of first-order quantification. For example, Kanellakis showed (cf. [AF90]) that there is a monadic NP sentence $\varphi(E, s, t)$ that says that there is an undirected path from s to t in the graph whose edge relation is given by E. Connectivity is then defined by the sentence $\forall x\, \forall y \varphi(E, x, y)$. However, as noted above, connectivity is not in monadic NP. So indeed, monadic NP is not closed under first-order quantification. Define *the positive first-order closure of monadic NP* to be the class that results from closing monadic NP under first-order quantification.[1] Thus, the positive first-order closure of monadic NP is the class of properties that are expressible by a sentence consisting of a string of

[1] In an earlier version of this paper, we referred to the positive first-order closure of monadic NP as "the first-order closure of monadic NP." What we call the first-order closure of monadic NP later in this paper was referred to in the earlier version of this paper as the "first-order/Boolean closure of monadic NP."

first-order quantifiers, followed by a string of existential unary second-order quantifiers, followed by a first-order formula. We use the word "positive," since the class is the result of closing monadic NP under the positive operations of first-order logic, namely first-order quantification, conjunction, and disjunction.[2] It is not hard to see that the positive first-order closure of monadic NP is a subclass of NP. As we noted above, there is an undirected graph property (namely, connectivity) that is in the positive first-order closure of monadic NP but not in monadic NP.

The main reason for our interest in monadic NP is that it is a tractable subclass of NP, in that we can prove lower bounds by showing that certain properties are not in monadic NP. This paper was motivated by the desire to develop new game-theoretic tools that will enable us to prove lower bounds involving extensions of monadic NP. We have mentioned one extension of monadic NP, namely, the positive first-order closure of monadic NP. We now give another. Define *the Boolean closure of monadic NP* to consist of those properties that are Boolean combinations of properties in monadic NP. If NP $\neq$ co-NP, then there is an undirected graph property (namely, non-3-colorability) that is in the Boolean closure of monadic NP but not in the positive first-order closure of monadic NP. The next theorem shows the converse (but without making any complexity-theoretic assumptions).

THEOREM 1.1. *There is an undirected graph property that is in the positive first-order closure of monadic NP but not in the Boolean closure of monadic NP.*

Define *closed monadic NP* to be the closure of monadic NP (or equivalently, of first-order logic) under first-order quantification and existential unary second-order quantification. Like before, closed monadic NP is a subclass of NP. We call this class simply "closed monadic NP" because it is formed by closing monadic NP under the two types of quantification that appear in the definition of monadic NP. Closed monadic NP differs from monadic NP in that we allow the possibility of arbitrary interleavings of first-order quantifiers among the existential unary second-order quantifiers. The next theorem shows that closed monadic NP contains a property that is not in either of the extensions of monadic NP that we just considered, namely the positive first-order closure of monadic NP and the Boolean closure of monadic NP. In fact, it shows even more. Define *the first-order closure of monadic NP* to be the result of closing monadic NP under both first-order quantification and Boolean operations. Thus, this class is the result of closing monadic NP under all of the first-order operations, namely first-order quantification, conjunction, disjunction, and negation.

THEOREM 1.2. *There is an undirected graph property that is in closed monadic NP but not in the first-order closure of monadic NP.*

[2] Conjunction and disjunction "come for free." For example, the conjunction of the sentences $\forall x\,\exists y\,\forall z\,\exists X\varphi$ and $\forall x'\,\exists y'\,\forall z'\,\exists X'\varphi'$, where the first-order variables x, y, z (resp., x', y', z') do not appear in φ' (resp., φ), and the unary second-order variable X (resp., X') does not appear in φ' (resp., φ), is equivalent to the sentence $\forall x\,\exists y\,\forall z\,\forall x'\,\exists y'\,\forall z'\,\exists X\,\exists X'(\varphi \wedge \varphi')$.

Closed monadic NP is a rich subclass of NP. Thus, it is a proper extension of the positive first-order closure of monadic NP (this follows from Theorem 1.2). Furthermore, the positive first-order closure of monadic NP is itself rich, in that it not only contains properties not in monadic NP, but even contains properties not in the Boolean closure of monadic NP (Theorem 1.1). We first realized the importance and naturalness of closed monadic NP because of the fact that connectivity is in this class (as the sentence $\forall x \, \forall y \varphi(E, x, y)$ discussed above shows). We have come up with a number of candidates for graph properties that we originally thought might not be in closed monadic NP, but which we showed, to our surprise, are indeed in this class. For example, we originally believed that (s, t)-connectivity of directed graphs, a property known not to belong to monadic NP [AF90], was a good candidate for a property not in closed monadic NP. However, we later showed that (s, t)-connectivity of directed graphs is in closed monadic NP (Theorem 4.1).

On the one hand, closed monadic NP is a monadic subclass of NP (perhaps even the "right" monadic subclass of NP), if we regard only the pattern of higher-order quantifiers (and thus ignore the first-order quantifiers). Indeed, there are often situations in mathematics where it is the higher-order quantifiers that have the important effect. For example, in the analytical hierarchy, it is only the quantifiers over the real numbers that matter: the quantifiers over natural numbers can be "swallowed up". On the other hand, when we alternate first-order quantifiers and existential unary second-order quantifiers, this behaves in some ways like existential second-order quantifiers of higher arity. Intuitively, for example, we can simulate $\forall x \, \exists R$, where x is a first-order variable and R represents a unary relation, by $\exists S \, \forall x$, where S is binary and where the unary relation R that corresponds to x is the set of all y such that Sxy. So $\forall x \, \exists R$ is behaving like a limited form of the existential binary second-order quantifier $\exists S$. This is significant, since it would be a huge step forward to extend our game-theoretic machinery to allow us to prove lower bounds about existential higher-arity quantification.[3] When we prove that some property is not in, for example, the positive first-order closure of monadic NP (such a result follows from Theorem 1.2), we are proving a lower bound about a limited form of existential higher-arity quantification.

To separate monadic NP from monadic co-NP, Fagin [Fag75] extended the theory of Ehrenfeucht–Fraïssé games to monadic NP games. In the standard (first-order) Ehrenfeucht–Fraïssé game over a pair G_0, G_1 of structures, two players, the spoiler and the duplicator, take turns placing pebbles on elements of the structures. Informally, the duplicator wins the game if the two structures, when restricted to the pebbled elements, are isomorphic (a more precise definition is given in Section 5). In this game (and in the other games we shall discuss), the way that we prove a lower bound, that is, an inexpressibility result, is to show that the duplicator has a winning strategy. In the monadic NP game, the spoiler starts by coloring the elements of G_0, the duplicator responds by coloring the elements of

[3] There has been some limited work on binary NP, which allows existential binary second-order quantifiers. Specifically, Durand *et al.* [DLS98] have shown results such as that existentially quantifying over unary functions is strictly less powerful than existentially quantifying over arbitrary binary relations.

G_1, and the two players then play the first-order game on the colored structures. Fagin showed that connectivity is not in monadic NP by showing that for every choice of the number of colors and the number of pebbling rounds, the duplicator has a winning strategy in the monadic NP game played over a pair G_0, G_1 of graphs, where G_0 consists of a single cycle and G_1 consists of two cycles. Since also, as we noted, connectivity is in monadic co-NP, this gives us the separation of monadic NP from monadic co-NP. Fagin *et al.* [FSV95] found a simpler proof that connectivity is not in monadic NP. Their proof makes use of several game-theoretic tools. They discuss the importance of developing a game-theoretic toolkit to help build toward our goal of applying game-theoretic techniques to fundamental problems such as whether NP equals co-NP.

In the game corresponding to the Boolean closure of monadic NP, played over graphs G_0 and G_1, the spoiler gets to choose which of G_0 and G_1 he wishes to color. After coloring this structure, the duplicator then colors the other structure. The two players then play the first-order game on the colored structures. In order to prove our result (Theorem 1.1) that there is an undirected graph property in the positive first-order closure of monadic NP but not in the Boolean closure of monadic NP, we must show the existence of a winning strategy for the duplicator in the game corresponding to the Boolean closure of monadic NP. Proving such a result is an advance in our goal of developing new game-theoretic techniques. This is because, apparently for the first time, we are being forced to consider a game where the spoiler has a choice of which structure to color.

In order to deal with the complexity introduced by the fact that the spoiler can choose which structure to color, we make use of two techniques. First, we create "superstructures," which are essentially structures of structures. For example, we create cycles, each of whose "points" is either a cycle or a disjoint union of cycles. Second, in considering the duplicator's strategy, we "split up" the (super)structures into disjoint pieces in different ways, where the way we split depends on which structure the spoiler colors. We then consider subgames that are applied separately to each of the disjoint pieces. We feel that these two techniques give us an approach that is both elegant and powerful.

Our most difficult result (Theorem 1.2) is the fact that there is an undirected graph property that is in closed monadic NP but is not in the first-order closure of monadic NP. In the game corresponding to the first-order closure of monadic NP, played over graphs G_0 and G_1, the spoiler not only gets to choose which of G_0 and G_1 he wishes to color, but he does not have to make his selection until after a number of pebbling moves have been played. Thus, not only are we faced with a situation where the spoiler gets to choose which structure to color, but apparently also for the first time, we are being forced to consider a game where there are pebbling rounds both before and after the coloring round. In order to prove that there is an undirected graph property that is in closed monadic NP but not in the first-order closure of monadic NP, we must show the existence of a winning strategy for the duplicator in this complicated game. To prove this result, we make use of one further technique. We create "super-superstructures" that are essentially structures of superstructures (and hence structures of structures of structures). Here each superstructure in the super-superstructure is one of the superstructures that we used in the previous proof that a certain undirected graph property is not in the Boolean closure of monadic NP.

We note that Ajtai and Fagin [AF90] have introduced a variation of the monadic NP game for which it is much easier to show that the duplicator has a winning strategy. In such games, the spoiler is forced to color G_0 without knowing what G_1 is. Unfortunately, for the games we use to prove the lower bound parts of Theorems 1.1 and 1.2, we cannot apply the idea behind the Ajtai–Fagin game to simplify the proofs that the duplicator has a winning strategy: the spoiler must know what G_0 and G_1 are before the coloring round.

We now discuss the relationship between closed monadic NP and the "monadic (polynomial-time) hierarchy." Monadic second-order logic allows both existential and universal second-order unary quantifiers. A property is said to be within the kth level of the monadic hierarchy if it is definable by a sentence of monadic second-order logic where all of the second-order quantifiers are at the beginning, and there are at most $k-1$ alternations between existential and universal second-order quantifiers. In particular, members of monadic NP are within the first level of the monadic hierarchy. It is easy to see that members of closed monadic NP are in the monadic hierarchy (we simply simulate first-order quantifiers by unary second-order quantifiers). Matz and Thomas [MT97b] have shown that the monadic hierarchy is strict. (We note that their proof is not game-theoretic.) A part of their proof is to define the class of "$\Sigma_1^{TC(1)}$ sentences" (essentially, monadic NP sentences where the first-order part can contain a transitive closure operator) and show that for each k, there is a property that is expressible by a $\Sigma_1^{TC(1)}$ sentence but is not within the kth level of the monadic hierarchy. We show that every property expressible by a $\Sigma_1^{TC(1)}$ sentence is in closed monadic NP, which gives us the following.

THEOREM 1.3. *For every k, there is an undirected graph property that is in closed monadic NP but not within the kth level of the monadic hierarchy.*

(This result was shown independently by Matz and Thomas [MT97a]. See [MST] for this and more recent improvements.) This is another indication of the richness of closed monadic NP.

We define the *closed monadic hierarchy* similarly to the monadic hierarchy, except that we focus only on the pattern of the higher-order quantifiers, and thus allow arbitrary interleavings of first-order quantifiers among the unary second-order quantifiers. We show that if the polynomial-time hierarchy is strict, then the closed monadic hierarchy is strict. It is an interesting open question as to whether we can prove the strictness of the closed monadic hierarchy without making any complexity-theoretic assumptions.

Proving the lower bounds in this paper that involve various extensions of monadic NP has required us to develop powerful new tools for the game-theoretic toolbox. Extending our results still further, such as proving that the closed monadic hierarchy is strict or proving that there is a property in the monadic hierarchy but not in closed monadic NP, will probably require further tools to be developed.

We now outline the remainder of the paper. Section 2 contains definitions of some basic concepts from logic and graph theory that are used in the paper. In this section we also review the definition of monadic NP. In Section 3 we give definitions of several extensions of monadic NP, such as closed monadic NP, that are the principal subjects of this paper, and we note some inclusions among these new classes. Section 4 contains examples showing that certain graph properties belong to certain of the descriptive complexity classes defined in the previous two sections. This section in particular contains our result that directed (s, t)-connectivity belongs to closed monadic NP. Many of the results in Section 4 are used as building blocks later in the paper. Section 5 reviews the first-order Ehrenfeucht–Fraïssé game and its extension to the monadic NP game. Section 6 recalls a theorem from [FSV95], based on a technique of Hanf, that gives a sufficient condition for the duplicator to win the first-order Ehrenfeucht–Fraïssé game; this result is used several times in the sequel. In Section 7 we prove a result that is needed later, and that is interesting in its own right: We show that for every choice of the number of colors and the number of pebbling rounds, the duplicator has a winning strategy in the monadic NP game played over a pair G_0, G_1 of graphs, where G_0 is a cycle C and G_1 consists of two disjoint copies of C. In [Fag75] a slightly weaker version of this theorem is given, where one of the cycles in G_1 is a cycle C' having size different than C. We need the stronger result for our purposes. Section 8 contains the proof of Theorem 1.1. In this section we also define a game that characterizes the Boolean closure of monadic NP and that we use to prove Theorem 1.1. Section 9 contains the proof of Theorem 1.2, as well as the definition of a game that characterizes the first-order closure of monadic NP. In Section 10 we review the definition of the monadic hierarchy, recall some results of Matz, Schweikardt, and Thomas [MT97b, Sch97, MST], and prove Theorem 1.3. In Section 11 we define the closed monadic hierarchy and show that the closed monadic hierarchy is strict if the polynomial-time hierarchy is strict. This follows from another result that we prove in Section 11, that there is an undirected graph property, an extension of 3-colorability, that belongs to the kth level of the monadic hierarchy and is complete for the class Σ_k^{P} in the polynomial-time hierarchy; this may be of independent interest as a new Σ_k^{P}-complete problem. In Section 12 we review the "growth" method that Matz and Thomas used to prove strictness of the monadic hierarchy (this method is not game-theoretic), and we give an alternate proof of Theorem 1.2 using a combination of a growth argument and a game argument. Section 13 contains conclusions and open questions.

2. DEFINITIONS AND CONVENTIONS

A *language* $\mathscr{L}$ (sometimes called a *similarity type*, a *signature*, or a *vocabulary*) is a finite set $\{P_1, ..., P_s\}$ of relation symbols, each of which has an arity.

An $\mathscr{L}$-*structure* (or *structure over* $\mathscr{L}$, or simply *structure*) is a set U (called the *universe*), along with a mapping associating a relation R_i over U with each $P_i \in \mathscr{L}$, where R_i has the same arity as P_i, for $1 \leqslant i \leqslant s$. We may call R_i the *interpretation of* P_i. By abuse of notation, we often let a relation symbol P also denote an interpretation of P. In addition, since we are primarily interested in the case that all quantified relation symbols are unary, for an interpretation of a unary relation symbol P we often let P also denote the set of points true under the interpretation, writing $x \in P$ synonymously with Px. The structure is called *finite* if the universe U

is. Unless otherwise stated, throughout the rest of this paper we make the assumption that all structures we consider are finite. Thus, this is a paper in finite model theory.

In this paper, we are especially interested in *graphs* and *colored graphs*. Graphs are simply structures where the language consists of a single binary relation symbol. Such structures are in general directed graphs. We often let such a structure represent an undirected graph by ignoring the directions of the edges. Thus, if E is a binary relation symbol, then the undirected graph that corresponds to an $\{E\}$-structure has the (symmetric) edge relation E' given by $E'xy = Exy \vee Eyx$. When talking about undirected graphs in the text, we say "there is an edge between x and y" or "x has an edge with y" to mean $E'xy$. In the context of directed graphs, the direction of the edge is made explicit, as in "there is an edge from x to y". We let (x, y) denote the undirected edge between x and y, and $\langle x, y \rangle$ denote the edge directed from x to y. It is convenient to assume that graphs have no self-loops, although this assumption does not affect our results. Colored graphs are structures where the language consists of a single binary relation symbol and some number of unary relation symbols. If G is a colored graph, where the interpretations of the unary relation symbols in the language are $R_1, \ldots, R_k$, then by the *color* of a point a in the universe of G, we mean a description of which R_i's the point a is a member of. Thus, intuitively, there are 2^k possible colors.

For definitions of a first-order formula (where, intuitively, the only quantification is over members of the universe, and not over, say, sets of members of the universe), and what it means for a structure S to *satisfy* a sentence σ, written $S \models \sigma$, see Enderton [End72] or Shoenfield [Sho67]. A *sentence* is a formula with no free (first-order) variables. We note that equality is treated as a special relation symbol, which is not considered to be a member of the language $\mathscr{L}$, and which always has the standard interpretation. The *quantifier depth* $QD(\varphi)$ of a first-order formula φ is defined recursively as follows: $QD(\varphi) = 0$ if φ is quantifier-free; $QD(\neg\varphi) = QD(\varphi)$; $QD(\varphi_1 \wedge \varphi_2) = \max\{QD(\varphi_1), QD(\varphi_2)\}$; $QD(\exists\varphi) = 1 + QD(\varphi)$.

In passing from first-order logic to second-order logic, we allow quantification over relations. In particular, a Σ_1^1 formula is a formula of the form $\exists A_1 \cdots \exists A_k \varphi$, where φ is first-order and where the A_i's are relation symbols. As an example, we now construct a Σ_1^1 sentence that says that a graph (with edge relation denoted by E) is 3-colorable. As before, let $E'xy$ denote $Exy \vee Eyx$. In this sentence, the three colors are represented by the unary relation symbols A_1, A_2, and A_3. Let φ_1 say "Each point has exactly one color". Thus, φ_1 is

$$\forall x((A_1x \wedge \neg A_2x \wedge \neg A_3x) \vee (\neg A_1x \wedge A_2x \wedge \neg A_3x) \vee (\neg A_1x \wedge \neg A_2x \wedge A_3x)).$$

Let φ_2 say "No two points with the same color have an edge between them". Thus, φ_2 is

$$\forall x \forall y((A_1x \wedge A_1y \Rightarrow \neg E'xy) \wedge (A_2x \wedge A_2y \Rightarrow \neg E'xy) \wedge (A_3x \wedge A_3y \Rightarrow \neg E'xy)).$$

The Σ_1^1 sentence $\exists A_1 \exists A_2 \exists A_3(\varphi_1 \wedge \varphi_2)$ then says "The graph is 3-colorable".

A Π_1^1 formula is a formula of the form $\forall A_1 \cdots \forall A_k \varphi$, where φ is first-order and where the A_i's are relation symbols. For example, letting φ_1 and φ_2 be as above, the Π_1^1 sentence $\forall A_1 \forall A_2 \forall A_3(\neg\varphi_1 \vee \neg\varphi_2)$ says "The graph is not 3-colorable".

A Σ_1^1 formula $\exists A_1 \cdots \exists A_k \varphi$, where φ is first-order, is said to be *monadic* if each of the A_i's is unary, that is, the existential second-order quantifiers quantify only over sets. A *monadic* Π_1^1 *formula* is defined similarly, restricting all (here universally) quantified relations to be unary. Following [FSV95], we often refer to a monadic Σ_1^1 formula (resp., monadic Π_1^1 formula) as a *monadic NP formula* (resp., *monadic co-NP formula*).

A *property* is a class of structures, all having the same signature, and closed under isomorphism. For example, the property "3-colorability" is the class of all $\{E\}$-structures corresponding to 3-colorable graphs. A sentence σ *expresses* the property $\mathscr{S}$ (having signature $\mathscr{L}$) if all $\mathscr{L}$-structures in the class $\mathscr{S}$ satisfy σ, and all $\mathscr{L}$-structures not in the class $\mathscr{S}$ do not satisfy σ. Define *monadic NP* (resp., *monadic co-NP*) to be the class of properties expressible by a monadic NP sentence (resp., monadic co-NP sentence). For example, the examples above show that 3-colorability is in monadic NP, and that non-3-colorability is in monadic co-NP, since the relations A_1, A_2, A_3 in these examples are unary.

To close this section, we define some terms concerning graphs. Additional graph terminology will be introduced later as needed. If G is a graph and X is a subset of the points of G, let $ind(X)$ denote the subgraph of G induced by X (*ind* abbreviates "induced"). A *connected component* of a graph G is an induced subgraph H such that H is connected, but such that no induced subgraph H' of G that is connected has H as a proper induced subgraph. If x is a point of the graph G, let $CC(x)$ denote the connected component of G that contains x. A *path* in an undirected graph is a sequence $x_0, x_1, ..., x_k$ of points, where $k \geqslant 0$, such that there is an edge between x_i and x_{i+1} for all $0 \leqslant i < k$; such a path is said to be a *path between* x_0 *and* x_k. We say that a path $x_0, ..., x_k$ is *strict* if there is no edge between x_i and x_j for all i, j with $|i-j| > 1$. If s and t are points of an undirected graph G, we say that s and t are *path-connected* if there is a path between s and t (equivalently, if s and t belong to the same connected component of G). If G is a directed graph with points s and t, we say that a sequence $x_0, ..., x_k$ of points, where $k \geqslant 0$, is a *directed path from s to t* if $s = x_0$, $t = x_k$, and there is an edge directed from x_i to x_{i+1} for all $0 \leqslant i < k$. In using these terms, the graph G will always be clear from context.

3. BEYOND MONADIC NP

We now formally define several types of formulas of interest to us. In parallel, for each type T of formula being defined, we define a corresponding descriptive complexity class containing those properties expressible by a sentence of type T. In our notations for formulas, MNP abbreviates monadic NP, FO abbreviates first-order, PFO abbreviates positive first-order, and BOOL abbreviates Boolean.

• A *BOOL(MNP) formula* is a Boolean combination of monadic NP formulas. The *Boolean closure of monadic NP* is the class of properties expressible by a BOOL(MNP) sentence.

• A *PFO(MNP) formula* is a formula of the form $\mathbf{P}\psi$, where $\mathbf{P}$ consists of first-order quantifiers, and ψ is a monadic NP formula. The *positive first-order closure of monadic NP* is the class of properties expressible by a PFO(MNP) sentence.

• A *FO(MNP) formula* is a formula in the class of formulas obtained by closing the monadic NP formulas under first-order quantification and Boolean operations. The *first-order closure of monadic NP* is the class of properties expressible by a FO(MNP) sentence.

• A *closed monadic NP formula* is a formula of the form $\mathbf{Q}\varphi$, where the prefix $\mathbf{Q}$ can have an arbitrary interleaving of first-order quantifiers and existential unary second-order quantifiers, and where φ is first-order. For example, $\mathbf{Q}$ could be $\forall x_1 \exists A_1 \forall x_2 \exists x_3 \exists A_2$, where $\forall x_1$, $\forall x_2$, and $\exists x_3$ are first-order quantifiers, and $\exists A_1$ and $\exists A_2$ are existential unary second-order quantifiers. *Closed monadic NP* is the class of properties expressible by a closed monadic NP sentence.

If ψ_1 and ψ_2 are closed monadic NP formulas, X denotes a unary relation, and x denotes a first-order variable, then $\exists X\psi_1$, $\exists x\psi_1$, and $\forall x\psi_1$ are closed monadic NP formulas, and $(\psi_1 \vee \psi_2)$ and $(\psi_1 \wedge \psi_2)$ are equivalent to closed monadic NP formulas. We often make use of these facts in constructing closed monadic NP formulas.

We now note some simple relationships among these descriptive complexity classes and standard computational complexity classes. First,

$$\begin{aligned}\text{monadic NP} &\subseteq \text{the positive first-order closure of monadic NP}\\ &\subseteq \text{closed monadic NP}\\ &\subseteq \text{NP}.\end{aligned}$$

The first two inclusions are obvious from the definitions. It is easy to establish the third inclusion by describing, for each closed monadic NP sentence σ, a nondeterministic polynomial-time Turing machine M that accepts precisely the structures that satisfy σ (where each structure is encoded as a string of symbols suitable as input to the Turing machine). We now sketch how M operates. Let σ have the form $\mathbf{Q}\varphi$, where $\mathbf{Q}$ contains an arbitrary interleaving of first-order quantifiers and existential unary second-order quantifiers, and where φ is first-order. The machine processes the quantifiers of $\mathbf{Q}$ in left-to-right order. For each existential unary second-order quantifier, $\exists A$, encountered, M nondeterministically guesses an interpretation of A. For each first-order quantifier, $\exists x$ or $\forall x$, encountered, M deterministically cycles through all assignments of elements of the universe to x. Given interpretations of all relations and assignments to all variables occurring free in φ, the truth of φ can be evaluated in polynomial time (as is well known). Since the number of first-order quantifiers in $\mathbf{Q}$ is fixed, the amount of time that M spends cycling through all assignments to the quantified variables is polynomial in the size of the universe.

As we now discuss, all of these inclusions are proper, even when we restrict our attention to undirected graph properties. As noted in the introduction, the first inclusion is known to be proper, since connectivity belongs to the positive first-order closure of monadic NP but does not belong to monadic NP. In fact, we actually prove the stronger result that there is an undirected graph property that is in the positive first-order closure of monadic NP but not in the Boolean closure of monadic NP (Theorem 1.1). As for the second inclusion, we actually prove the stronger result that there is an undirected graph property that is in closed monadic NP but not in the first-order closure of monadic NP (Theorem 1.2). The third inclusion is proper since Turán [Tur84] has shown that the "perfect matching" property, the class of undirected graphs having a perfect matching, cannot be expressed in monadic second-order logic. (As with Matz and Thomas' result, the proof is not game-theoretic.) It is well known that perfect matching belongs to P, and clearly every closed monadic NP sentence is a sentence of monadic second-order logic. We note that Turán also showed that Hamiltonicity cannot be expressed in monadic second-order logic.

For classes involving Boolean combinations, we have

$$\begin{aligned}\text{monadic NP} \cup \text{monadic co-NP} &\subseteq \text{the Boolean closure of monadic NP}\\ &\subseteq \text{the first-order closure of monadic NP}\\ &\subseteq \mathrm{P}^{\mathrm{NP}}.\end{aligned}$$

The first two inclusions are obvious from the definitions. We now show the third inclusion. By definition, a property is in the class P^{NP} if it is recognized by some deterministic polynomial-time oracle machine with an NP oracle. Let σ be a FO(MNP) sentence. Thus, σ is constructed by applying first-order quantification and Boolean operations to a set S of monadic NP formulas. To determine if a given structure satisfies σ, each first-order quantifier of the form $\exists x$ (resp., $\forall x$) that was used to construct σ from S is replaced by a disjunction (resp., conjunction) over all x in the universe. In this way, σ is transformed in polynomial time (i.e., polynomial in the size of the universe) to a Boolean combination of monadic NP sentences (each such sentence is a formula in S with each of its free first-order variables replaced by an element of the universe). Finally, the truth of each of these monadic NP sentences can be determined using an NP oracle (recall that monadic NP is contained in NP). This actually shows the stronger result that the first-order closure of monadic NP is contained in $\mathrm{P}_{\|}^{\mathrm{NP}}$, defined as the class of properties that are recognized by some deterministic polynomial-time oracle machine with an NP oracle where all of the oracle queries must be made in parallel (unlike P^{NP} where each query can depend on the answers to previous queries); for definitions and results concerning $\mathrm{P}_{\|}^{\mathrm{NP}}$ see, for example, [Wag90]. In fact, by combining a syntactic characterization of P^{NP} given by Buss and Hay [BH91] with the fact that $\Sigma_1^1 = \mathrm{NP}$, it is not hard to show that the properties in $\mathrm{P}_{\|}^{\mathrm{NP}}$ are precisely the properties in the first-order closure of Σ_1^1, provided that Σ_1^1 formulas can use a built-in linear order relation. In other words, if we extend the first-order closure of monadic NP in two ways, first by extending monadic NP to (non-monadic) Σ_1^1, and second by including a built-in linear order, then we get precisely $\mathrm{P}_{\|}^{\mathrm{NP}}$.

Once again, as we now discuss, all of these inclusions are proper, even when we restrict our attention to undirected graph properties. The properness of the first inclusion follows from the fairly simple-to-show fact (Theorem 8.1) that the property "There are exactly two connected components" belongs to the Boolean closure of monadic NP but does not belong to monadic NP nor to monadic co-NP. As for the second inclusion, we actually prove the stronger result that there is an undirected graph property that is in the positive first-order closure of monadic NP but not in the Boolean closure of monadic NP (Theorem 1.1). Properness of the third inclusion is shown by Turán's result mentioned above. Alternatively, the properness of the third inclusion follows from the result that there is an undirected graph property that is in closed monadic NP (and hence in NP, and so in P^{NP}), but not in the first-order closure of monadic NP (Theorem 1.2).

Another inclusion, not shown above, is that the Boolean closure of monadic NP is contained in the Boolean hierarchy [CGH+88], the Boolean closure of NP; this inclusion is also proper, as shown by Turán's result.

4. EXAMPLES

This section contains several examples showing how various graph connectivity and nonconnectivity properties can be expressed by the types of sentences defined in the previous sections. Many of these results will be used as building blocks later in the paper. A new result is that the property of directed graphs with distinguished points s and t, that "There is a directed path from s to t", belongs to closed monadic NP. Ajtai and Fagin [AF90] call this property "directed reachability". They prove that this property does not belong to monadic NP.

We begin by showing that the "nonconnectivity" property, the class of nonconnected undirected graphs, belongs to monadic NP (this demonstration is from [Fag75]). Let φ_1 say "The set A is nonempty and its complement is nonempty", that is, $\exists x\, \exists y(Ax \wedge \neg Ay)$. Let φ_2 say "There is no edge between A and its complement", that is, $\forall x\, \forall y((Ax \wedge \neg Ay) \Rightarrow \neg E'xy)$. It is clear that the monadic NP sentence $\exists A(\varphi_1 \wedge \varphi_2)$ expresses nonconnectivity. Having described sentences for 3-colorability and nonconnectivity in detail, in the sequel we do not describe the first-order part of sentences in complete detail, but rather give enough of a high-level description that translation to formal logic is straightforward.

Similarly, the property "*ind*(X) is nonconnected" belongs to monadic NP (viewed as a property of a structure with signature containing a binary edge relation and a unary relation X). This is expressed by saying that there exists a set A of points such that $A \cap X$ and $\bar{A} \cap X$ are both nonempty (where $\bar{A}$ denotes the complement of A), and there is no edge between a point of $A \cap X$ and a point of $\bar{A} \cap X$. This is an example of a general principle, used later, that if a property $\mathscr{S}$ can be expressed by a monadic NP sentence (resp., closed monadic NP sentence) σ, then the property "*ind*(X) has property $\mathscr{S}$" can also be expressed by a monadic NP sentence (resp., closed monadic NP sentence). The new sentence is obtained from σ essentially by restricting everything to points in X.

We next recall the fact, due to Kanellakis (cf. [AF90]) that for undirected graphs, "s and t are path-connected" belongs to monadic NP. This is expressed by saying that either $s = t$, or there exists a set A of points such that the following conjunction holds: (a) s and t both belong to A; (b) s has an edge with exactly one point of A; (c) t has an edge with exactly one point of A; and (d) every point of A

except s and t has an edge with exactly two points of A. To see that this expresses that s and t are path-connected for $s \neq t$, first note that if s and t are path-connected, we can choose A to be the points on a shortest path between s and t. Second, if A satisfies (a)–(d), then, since the graph is finite, by "following" the path A from s we must eventually reach t. (The second argument fails if G is infinite. In fact, it is not hard to show by a compactness argument that "s and t are path-connected" cannot be expressed by a monadic NP sentence in the infinite case.) For future use, let the first-order formula $path(A, s, t)$ say that either (i) $s = t$ and $A = \{s\}$, or (ii) the conjunction (a)–(d) holds. That is, $path(A, s, t)$ says that (some sequential ordering of the points in) A is a strict path between s and t.

As we have noted, connectivity does not belong to monadic NP. However, the formula *path* can be used to show that connectivity belongs to the positive first-order closure of monadic NP, since the sentence $\forall x\, \forall y\, \exists A(path(A, x, y))$ expresses connectivity.

We now consider the directed reachability property mentioned at the beginning of this section. Ajtai and Fagin [AF90] show that this property does not belong to monadic NP. However, this property can be expressed by a closed monadic NP sentence, allowing first-order quantifiers to be interleaved with existential unary second-order quantifiers, as we now show.

THEOREM 4.1. *The directed graph property "There is a directed path from s to t" belongs to closed monadic NP.*

Proof. We show that the following closed monadic NP sentence expresses directed reachability, where A, P, S denote sets of points (unary relation symbols) and x, y denote points (variables symbols):

$(\exists A \text{ with } s \in A)(\forall x \in A \text{ with } x \neq t)(\exists y \in A)(\exists P)(\exists S)$

(1) P and S partition A, i.e., $P \cup S = A$ and $P \cap S = \varnothing$;

(2) $s, x \in P$ and $y \in S$;

(3) there is an edge from x to y; and

(4) the only edge from a point of P to a point of S is the edge from x to y.

To see that this expresses that "There is a directed path from s to t", say first that there is such a path. We show that the sentence is satisfied. Let $s = x_0, x_1, ..., x_k = t$ be a shortest directed path from s to t. Choose $A = \{x_0, x_1, ..., x_k\}$. Given an arbitrary $x \in A$ with $x \neq t$, let j be such that $x = x_j$ (note that $j < k$), and choose $y = x_{j+1}$, $P = \{x_i \mid 0 \leqslant i \leqslant j\}$, and $S = \{x_i \mid j+1 \leqslant i \leqslant k\}$. (Think of P as the set of predecessors of x (including x itself) in the directed path, and S as the set of successors of x.) Since $x_0, x_1, ..., x_k$ is a shortest directed path from s to t, it is easy to see that (1)–(4) are all satisfied. In particular, (4) is satisfied because, for all i, j with $0 \leqslant i < j \leqslant k$, there is an edge from x_i to x_j iff $j = i + 1$.

Say now that the sentence is satisfied by a directed graph G with distinguished points s and t. We show that there is a directed path from s to t. Using that the sentence is satisfied we show how, starting at s, to extend the path one point at a time. Let the set A be such that the part of the sentence following $\exists A$ is satisfied. Say that a sequence $x_0, x_1, ..., x_j$ of points with $j \geqslant 0$ is a *partial path* if the points all belong to A, the points are all distinct, $x_0 = s$, and for all i with $0 \leqslant i < j$ there is an edge from x_i to x_{i+1}. We show that if $x_0, x_1, ..., x_j$ is a partial path with $x_j \neq t$, then there is a point x_{j+1} such that $x_0, x_1, ..., x_{j+1}$ is a partial path. Since the singleton sequence containing only s is a partial path and since the graph is finite, this suffices to show that there is a directed path from s to t. Let $s = x_0, x_1, ..., x_j$ be a partial path with $x_j \neq t$. Instantiating $x = x_j$ in the sentence, let y, P, S be such that $y \in A$ and (1)–(4) are satisfied. Define $x_{j+1} = y$. By (3), there is an edge from $x = x_j$ to $y = x_{j+1}$. To complete the proof that $x_0, x_1, ..., x_{j+1}$ is a partial path, we must show that $x_i \neq x_{j+1}$ for $0 \leqslant i \leqslant j$. Suppose for contradiction that $x_i = x_{j+1}$ for some i with $0 \leqslant i \leqslant j$. Since $x_i = x_{j+1} = y \in S$, we have $x_i \in S$. Let k be the smallest integer such that $x_k \in S$. Since $x_0 = s \in P$ and $x_i \in S$, we have $1 \leqslant k \leqslant i$. Therefore, $k - 1 \geqslant 0$, and $x_{k-1} \in P$ by the choice of k. Recalling that the points $x_0, x_1, ..., x_j$ are distinct and that $k \leqslant i \leqslant j$, we see that the edge $\langle x_{k-1}, x_k \rangle$ is different from the edge $\langle x, y \rangle = \langle x_j, x_i \rangle$ because $x_{k-1} \neq x_j$. This contradicts the requirement (4) that the only edge from a point of P to a point of S is the edge $\langle x, y \rangle$. ∎

Noting that S is equivalent to $A \wedge \neg P$, the proof can easily be modified to show that directed reachability can be expressed by a sentence of the form $\exists A\, \forall x\, \exists P \varphi$, where A and P are unary relation symbols, x is a first-order variable, and φ is first-order.

Marcinkowsky [Mar99] has shown that directed reachability is not in the positive first-order closure of monadic NP, which answers a question that was left open in a previous version of our paper. It follows that a quantifier prefix of the form $\exists A\, \forall x\, \exists P$ is the unique, minimum-length prefix form containing existential unary second-order quantifiers and arbitrary first-order quantifiers such that, when placed in front of some first-order formula, the resulting sentence expresses directed reachability.

5. EHRENFEUCHT–FRAÏSSÉ GAMES

Ehrenfeucht–Fraïssé games [Ehr61, Fra54] are tools that have been very useful in showing that certain properties cannot be expressed by certain types of logics. For an introduction to Ehrenfeucht–Fraïssé games and some of their applications to finite model theory, see [AF90, pp. 122–126].

We begin with an informal definition of an *r-round first-order Ehrenfeucht–Fraïssé game* (where r is a positive integer), which we shall call an *r-game* for short. It is straightforward to give a formal definition, but we shall not do so. For ease in description, we restrict our attention to colored graphs, but everything we say generalizes easily to arbitrary structures. There are two *players*, called *the spoiler* and *the duplicator*, and two colored graphs, G_0 and G_1. In the first round, the spoiler selects a point in one of the two colored graphs, and the duplicator selects a point in the other colored graph. Let a_1 be the point selected in G_0, and let b_1 be the point selected in G_1. Then the second round begins, and again, the spoiler selects a point in one of the two colored graphs, and the duplicator selects a point in the other colored graph. Let a_2 be the point selected in G_0, and let b_2 be the point selected in G_1. This continues for r rounds. The duplicator wins if the colored subgraph of G_0 induced by $\langle a_1, ..., a_r \rangle$ is isomorphic to the colored subgraph of G_1 induced by $\langle b_1, ..., b_r \rangle$, under the function that maps a_i onto b_i for $1 \leqslant i \leqslant r$. That is, for the duplicator to win, (a) $a_i = a_j$ iff $b_i = b_j$, for each i, j; (b) $\langle a_i, a_j \rangle$ is an edge in G_0 iff $\langle b_i, b_j \rangle$ is an edge in G_1, for each i, j; and (c) a_i has the same color as b_i, for each i. Otherwise, the spoiler wins. We say that the spoiler or the duplicator *has a winning strategy* if he can guarantee that he will win, no matter how the other player plays. Since the game is finite, and there are no ties, the spoiler has a winning strategy iff the duplicator does not. If the duplicator has a winning strategy, then we write $G_0 \sim_r G_1$. In this case, intuitively, G_0 and G_1 are indistinguishable by an r-game.

The following important theorem (from [Ehr61, Fra54]) shows why these games are of interest. If $\mathscr{S}$ is a class of colored graphs, then let $\bar{\mathscr{S}}$ be the complement of $\mathscr{S}$, that is, the class of colored graphs not in $\mathscr{S}$.

THEOREM 5.1. *$\mathscr{S}$ is first-order expressible iff there is r such that whenever $G_0 \in \mathscr{S}$ and $G_1 \in \bar{\mathscr{S}}$, then the spoiler has a winning strategy in the r-game over G_0, G_1.*

Fagin [Fag75] introduced an Ehrenfeucht–Fraïssé game corresponding to monadic NP. Let G_0, G_1 be structures (here, G_0 and G_1 are not colored), and let c, r be positive integers (where c represents the number of colors and r the number of rounds). We call this game the *asymmetric (c, r)-game over G_0, G_1*. The rules are as follows.

1. The spoiler colors G_0 with the c colors.
2. The duplicator colors G_1 with the c colors.
3. The spoiler and duplicator play an r-game on the colored G_0, G_1.

The winner is decided as before. Of course, the isomorphism must respect colors.

We call the game asymmetric, since unlike the first-order game, the rules are asymmetric in G_0, G_1, in that the spoiler must color G_0. We have the following theorem, analogous to Theorem 5.1.

THEOREM 5.2 [Fag75]. *$\mathscr{S}$ is in monadic NP iff there are c, r such that whenever $G_0 \in \mathscr{S}$ and $G_1 \in \bar{\mathscr{S}}$, then the spoiler has a winning strategy in the asymmetric (c, r)-game over G_0, G_1.*

Ajtai and Fagin [AF90] introduced a variation of the asymmetric (c, r)-game and showed that it also characterizes monadic NP. As in [FSV95] we call this variation the Ajtai–Fagin game. It is played over a class $\mathscr{S}$ of structures, rather than over a pair of structures. Note that we can view the asymmetric (c, r)-game as a game over a class $\mathscr{S}$: in the first move, the duplicator chooses $G_0 \in \mathscr{S}$ and $G_1 \notin \mathscr{S}$, and the game then proceeds as in the asymmetric (c, r)-game over G_0, G_1. The Ajtai–Fagin variation interchanges two of the moves. In the *Ajtai–Fagin (c, r)-game over $\mathscr{S}$*, the duplicator chooses $G_0 \in \mathscr{S}$, the spoiler colors G_0, the duplicator chooses $G_1 \notin \mathscr{S}$ and colors it, and the spoiler and duplicator play an r-game on the colored G_0, G_1; the winner is determined as before. The Ajtai–Fagin game is seemingly easier for the duplicator to win, because he can base his choice of G_1 on the spoiler's coloring of G_0. (See [Fag97] for a discussion of a precise sense in which the Ajtai–Fagin game is easier for the duplicator to win.) Nevertheless, Ajtai and Fagin [AF90] show that $\mathscr{S}$ is in monadic NP iff there are

c, r such that the spoiler has a winning strategy in the Ajtai–Fagin (c, r)-game over $\mathcal{S}$. Thus, using the Ajtai–Fagin game, it is potentially easier to show that certain properties are not in monadic NP. In fact, every paper on monadic NP except the first two ([Fag75] and [dR87]) uses the Ajtai–Fagin game rather than the asymmetric (c, r)-game to prove lower bounds. For example, using the Ajtai–Fagin game (together with another tool described in Section 6), [FSV95] gives a very simple proof that connectivity is not in monadic NP. Unfortunately, we cannot use the idea behind the Ajtai–Fagin game to simplify the Ehrenfeucht–Fraïssé games that we use to prove our main results. This occurs for two reasons. First, both of the games that we use have a *symmetric* coloring round, in that the spoiler can choose which one of G_0 or G_1 to color, and the duplicator must then color the other. Second, one of these games involves a first-order (point-choosing) game before the structures are colored. For both reasons, the spoiler must know what both structures G_0 and G_1 are before the spoiler's coloring step occurs.

6. HANF'S TECHNIQUE

In this section, we state a result from [FSV95] which provides a simple but very useful sufficient condition for guaranteeing that $A \sim_r B$ for two structures A, B. The proof is based on a technique of Hanf [Han65].

Let A be an $\mathcal{L}$-structure, where $\mathcal{L} = \{P_1, ..., P_s\}$, and where R_i is the interpretation in A of the relation symbol P_i, for $1 \leqslant i \leqslant s$. Let a and b be two points in (the universe of) A. We say that a and b are *adjacent* (in A) if either $a = b$ or there is some R_i and some tuple t such that $t \in R_i$ and such that a and b are entries in the tuple t. Intuitively, two points a and b are adjacent if they are either identical or directly related by some relation of A. The *degree* of a point a is the cardinality of the set of points adjacent to a but not equal to a.

Essentially following Hanf, we define the *neighborhood* $Nbd(d, a)$ *of radius d about a* recursively as follows:

$$Nbd(1, a) = \{a\}$$

$$Nbd(d+1, a) = \{x \mid x \text{ is adjacent to some } b \in Nbd(d, a)\}$$

It is helpful to think of these neighborhoods as *open* spheres. Thus, intuitively, $Nbd(d, a)$ consists of all points whose distance from a is *strictly* less than d. Note that because a is adjacent to itself, we have that $Nbd(d, a) \subseteq Nbd(d+1, a)$.

Also following Hanf, we define the *d-type* of a point a to be the isomorphism type of the neighborhood of radius d about a with a as a distinguished point. Thus, the points a in A and b in B have the same d-type precisely if the substructure of A induced by $Nbd(d, a)$ is isomorphic to the substructure of B induced by $Nbd(d, b)$, under an isomorphism mapping a to b.

Let d, m be positive integers. We say that $\mathcal{L}$-structures A and B are (d, m)-*equivalent* if for every d-type τ, either A and B have the same number of points with d-type τ, or else both have at least m points with d-type τ. Intuitively, A and B are (d, m)-equivalent if for every d-type τ, they have the same number of points with d-type τ, where we can count only as high as m.

The following result is proved in [FSV95], based on the proof of a similar result of Hanf [Han65]. This version of Hanf's result from [FSV95] is useful in the context of finite structures, whereas Hanf's original result is not.

THEOREM 6.1. *Let r, f be positive integers. There are positive integers d, m, where d depends only on r, such that whenever A and B are (d, m)-equivalent structures where every point has degree at most f, then $A \sim_r B$.*

We remark that the proof of this result in [FSV95] describes an explicit winning strategy for the duplicator.

7. ONE CYCLE VERSUS TWO CYCLES

A directed cycle of length n consists of points α_i for $0 \leqslant i < n$ and edges $\langle \alpha_i, \alpha_{i+1} \rangle$ for $0 \leqslant i < n$, where subscripts are reduced modulo n. Because of our convention that structures are directed, "cycle" will always mean "directed cycle" when used in the context of structures. A result that will be useful later in the paper is that, for every c, r, there is a cycle C such that the duplicator has a winning strategy in the asymmetric (c, r)-game over the structures C, $C \oplus C$ (where $A \oplus B$ is the disjoint union of the structures A and B, and where we view different occurrences of C as being structures with disjoint universes). Actually, this result is implicit in the proof of Theorem 3 of Fagin [Fag75]. However, Theorem 3 in [Fag75] is stated in a weaker form, from which we can conclude only that there are two cycles C and C', possibly of different lengths, such that the duplicator wins the asymmetric (c, r)-game over C, $C \oplus C'$. For completeness, we give here a proof of the stronger result we use. As an added benefit, the proof given here is considerably simpler than the proof in [Fag75], because we are able to use a tool, Hanf's technique (Theorem 6.1), which was developed since the appearance of [Fag75]. Whereas Fagin described the duplicator's pebbling strategy explicitly, we use Hanf's technique to show that a strategy exists.

THEOREM 7.1. *Let c, r be positive integers. There are arbitrarily large cycles C such that the duplicator has a winning strategy in the asymmetric (c, r)-game over C, $C \oplus C$.*

This theorem follows immediately from the next lemma by taking C to be an arbitrary cycle of length n where $n \geqslant t$ and z divides n.

LEMMA 7.2. *Let c, r be positive integers. There exist positive integers t, z, such that if C is a cycle of length n with $n \geqslant t$ and C' is a cycle of length divisible by z, then the duplicator has a winning strategy in the asymmetric (c, r)-game over C, $C \oplus C'$.*

Proof. Assume without loss of generality that $c \geqslant 2$. Given c, r, let d, m be the integers obtained from Theorem 6.1 for this r and $f = 2$ (since we are dealing only with cycles, all points have degree 2). Values for t and z will be determined as the proof proceeds.

It is convenient to give different names to different occurrences of cycles. Let A be a cycle and n its length. Let $\alpha_0, \alpha_1, ..., \alpha_{n-1}$ be the points in order around A, so there is an edge from α_i to α_{i+1} for $0 \leqslant i < n$; here and subsequently, subscripts are reduced modulo n to belong to the interval $[0, n-1]$. Suppose that the spoiler colors the points of A with c colors. Let $\chi(\alpha_i)$ denote the color of α_i, and let $\chi(A)$ denote the colored structure A. Assuming that $n \geqslant 2d$ (we require $t \geqslant 2d$), the d-type of the point α_i in $\chi(A)$ is fully described by the following vector of $2d-1$ colors:

$$\langle \chi(\alpha_{i-(d-1)}), ..., \chi(\alpha_{i-1}), \chi(\alpha_i), \chi(\alpha_{i+1}), ..., \chi(\alpha_{i+(d-1)}) \rangle. \tag{1}$$

Let τ_i denote the d-type of α_i.

Let $N = c^{2d-1}$ be the number of d-types. Say that d-type τ is *rare* if it occurs less than m times as a d-type of a point of $\chi(A)$. Say that τ is *frequent* if it is not rare. The total number of occurrences of rare d-types is at most $(m-1)N$. Since this is some constant depending on c and r, but not on n or the coloring, it follows, as we shall show, that for n sufficiently large, there must be two points α_p and α_q with $0 \leqslant p < q < n$ such that: (1) $\tau_p = \tau_q$; (2) τ_i is frequent for all i with $p \leqslant i \leqslant q$; (3) α_p and α_q are at least distance d apart, that is, $\alpha_p \notin Nbd(d, \alpha_q)$; and (4) $q - p \leqslant dN$. To see this, note that there is a t depending on d, N, and the number of rare d-types (i.e., depending on c and r), such that for all $n \geqslant t$ and all colorings of A with c colors, there is a point α_k such that τ_{k+i} is frequent for $0 \leqslant i \leqslant dN$, and α_k and α_{k+dN} are at least distance d apart. Among the $dN+1$ points α_{k+i} for $0 \leqslant i \leqslant dN$, some subset of $d+1$ points must have the same d-type. Two points in this subset must be at least distance d apart.

Choose z so that z is divisible by every integer from 2 to dN. Let B be a cycle of length n and B' be a cycle of length n' divisible by z. Note that $n' \geqslant 2d$ since $z \geqslant dN \geqslant 2d$. Let β_i for $0 \leqslant i < n$ (resp., β'_i for $0 \leqslant i < n'$) be the points in order around B (resp., B'). The duplicator's coloring of B is copied from A, that is, $\chi(\beta_i) = \chi(\alpha_i)$ for all i. The duplicator's coloring of B' is obtained by repeating the color sequence $\chi(\alpha_p), \chi(\alpha_{p+1}), ..., \chi(\alpha_{q-1})$ around B for $n'/(q-p)$ times. Since $q-p \leqslant dN$, the choice of z ensures that n' is divisible by $(q-p)$. More precisely, $\chi(\beta'_i) = \chi(\alpha_{p+j})$ where $j \equiv i \bmod (q-p)$ and $0 \leqslant j < q-p$. Let $\chi(B)$ and $\chi(B')$ denote the structures B and B', respectively, after being colored as just described. Let τ'_i be the d-type of β'_i in $\chi(B')$. Since $\tau_p = \tau_q$ and α_p and α_q are at least distance d apart, it is easy to see that the sequence $\tau_p, \tau_{p+1}, ..., \tau_{q-1}$ of d-types repeats around $\chi(B')$ in the same cyclic way that the colors repeat, that is, $\tau'_i = \tau_{p+j}$ where $j \equiv i \bmod (q-p)$ and $0 \leqslant j < q-p$. Since τ_i is frequent for $p \leqslant i \leqslant q$, all d-types occurring in $\chi(B')$ are frequent. It follows that $\chi(A)$ and $\chi(B) \oplus \chi(B')$ are (d, m)-equivalent, since each rare d-type occurs exactly the same number times in $\chi(B)$ as it does in $\chi(A)$ and it does not occur in $\chi(B')$ at all, and since each frequent d-type occurs at least m times in $\chi(B)$. It follows from Theorem 6.1 that $\chi(A) \sim_r \chi(B) \oplus \chi(B')$, so the duplicator wins. ▮

8. THE BOOLEAN CLOSURE OF MONADIC NP

In this section we investigate the expressive power of the Boolean closure of monadic NP. In addition to establishing results about the Boolean closure of monadic NP, the results and techniques in this section are useful in the next section on the power of closed monadic NP. The main result of this section establishes a limit on the power of the Boolean closure of monadic NP. We show that there is an undirected graph property in the positive first-order closure of monadic NP that is not in the Boolean closure of monadic NP. We actually show that this property is expressible by a sentence of the form $\forall x \forall y \psi$ where ψ is a monadic NP formula; so, in some sense, this property is not too far from monadic NP. Another (relatively easy) result of this section is that there is an undirected graph property that belongs to the Boolean closure of monadic NP, but does not belong to monadic NP nor to monadic co-NP. We begin with this result. Its proof uses, in a simple way, a technique of splitting structures into disjoint parts and describing different Ehrenfeucht–Fraïssé game strategies for different parts; this technique is used later in a more complicated way. Let the property "There are exactly two connected components" be the class of undirected graphs that have exactly two connected components.

THEOREM 8.1. *The property "There are exactly two connected components" belongs to the Boolean closure of monadic NP, but does not belong to monadic NP nor to monadic co-NP.*

Proof. Let $\mathcal{S}$ be the property "There are exactly two connected components". We now show that $\mathcal{S}$ belongs to the Boolean closure of monadic NP. We first observe that, for each $k \geqslant 2$, there is a monadic NP sentence γ_k that says that G has at least k connected components. The sentence γ_k is similar to the sentence described in Section 4 for nonconnectivity. The sentence γ_k says that there exist k nonempty sets $A_1, ..., A_k$, where the sets are pairwise disjoint, and such that there is no edge between a point of A_i and a point of $\overline{A_i}$ for $1 \leqslant i \leqslant k$. Thus, $\mathcal{S}$ is expressed by the sentence $\gamma_2 \wedge \neg\gamma_3$, a Boolean combination of monadic NP sentences.

To show that $\mathcal{S}$ does not belong to monadic co-NP, it suffices to show that the complementary property $\bar{\mathcal{S}}$ does not belong to monadic NP. Using Theorem 5.2, we need only show that for every c, r there are $G_0 \in \bar{\mathcal{S}}$ and $G_1 \in \mathcal{S}$ such that the duplicator has a winning strategy in the asymmetric (c, r)-game over G_0, G_1. Let c, r be given. By Theorem 7.1, there is a cycle C such that the duplicator has a winning strategy in the asymmetric (c, r)-game over C, $C \oplus C$. So we need only take $G_0 = C$ and $G_1 = C \oplus C$.

To show that $\mathcal{S}$ does not belong to monadic NP, we again use Theorem 5.2, the difference being that now we must choose $G_0 \in \mathcal{S}$ and $G_1 \in \bar{\mathcal{S}}$. For a given c, r, let C be as above. Now we take $G_0 = C \oplus C$ and $G_1 = C \oplus C \oplus C$. We split G_0 and G_1 into disjoint pieces as follows. Let $G'_0 = G''_0 = C$, $G'_1 = C$, and $G''_1 = C \oplus C$. To show that the duplicator has a winning strategy in the asymmetric (c, r)-game over G_0, G_1, it is clearly sufficient to show that (1) the duplicator has a winning strategy in the asymmetric (c, r)-game over G'_0, G'_1, and (2) the duplicator has a winning strategy in the asymmetric (c, r)-game over G''_0, G''_1. For (1), this is obvious because G'_0 and G'_1 are isomorphic. For (2), this follows by the choice of C. ∎

We now define a game that (as we shall show) corresponds to properties that are in the Boolean closure of monadic NP. Let G_0, G_1 be structures, and let c, r be positive integers (where c represents the number of colors and r the number of rounds). We call this game the *symmetric (c, r)-game over G_0, G_1*. The rules are as follows.

1. The spoiler colors either G_0 or G_1 with the c colors.
2. The duplicator colors the other structure with the c colors.
3. The spoiler and duplicator play an r-game on the colored G_0, G_1.

The winner is decided as before. Again, the isomorphism must respect colors.

The difference between the symmetric and asymmetric games is that in the asymmetric game, the spoiler must color G_0 and the duplicator G_1, whereas in the symmetric game, the spoiler can color either structure and the duplicator then colors the other structure. Clearly, the duplicator has a winning strategy in the symmetric (c, r)-game over G_0, G_1 iff (1) the duplicator has a winning strategy in the asymmetric (c, r)-game over G_0, G_1 and (2) the duplicator has a winning strategy in the asymmetric (c, r)-game over G_1, G_0. We have the following theorem, analogous to Theorem 5.2.

THEOREM 8.2. *$\mathcal{S}$ is in the Boolean closure of monadic NP iff there are c, r such that whenever $G_0 \in \mathcal{S}$ and $G_1 \in \bar{\mathcal{S}}$, then the spoiler has a winning strategy in the symmetric (c, r)-game over G_0, G_1.*

Proof. Assume first that $\mathcal{S}$ is in the Boolean closure of monadic NP. Thus, there is a sentence φ that is a Boolean combination of monadic NP sentences, such that every member of $\mathcal{S}$ satisfies φ, and no member of $\bar{\mathcal{S}}$ satisfies φ. By converting φ to disjunctive normal form, we can assume without loss of generality that φ is of the form $\bigvee_{i=1}^{m_1} \bigwedge_{j=1}^{m_2} \varphi_{i,j}$, where each $\varphi_{i,j}$ is a monadic NP sentence or a monadic co-NP sentence. Since m_1 and m_2 are finite, there are k, r such that every $\varphi_{i,j}$ is of the form $\exists A_1 \cdots \exists A_{k'} \psi$ (if $\varphi_{i,j}$ is monadic NP) or $\forall A_1 \cdots \forall A_{k'} \psi$ (if $\varphi_{i,j}$ is monadic co-NP), where the A_ℓ's are unary, where $k' \leqslant k$ and where ψ has quantifier depth at most r. Let $c = 2^k$. We now show that whenever $G_0 \in \mathcal{S}$ and $G_1 \in \bar{\mathcal{S}}$, then the spoiler has a winning strategy in the symmetric (c, r)-game over G_0, G_1.

Assume that $G_0 \in \mathcal{S}$ and $G_1 \in \bar{\mathcal{S}}$. By definition of φ, every member of $\mathcal{S}$ satisfies φ, and no member of $\bar{\mathcal{S}}$ satisfies φ. In particular, $G_0 \models \varphi$, and $G_1 \not\models \varphi$. Since $G_0 \models \varphi$, there exists i_0 such that $G_0 \models \bigwedge_{j=1}^{m_2} \varphi_{i_0,j}$. Since $G_1 \not\models \varphi$, it follows that $G_1 \not\models \bigwedge_{j=1}^{m_2} \varphi_{i_0,j}$. Therefore, there exists j_0 such that $G_1 \not\models \varphi_{i_0,j_0}$. Since $G_0 \models \bigwedge_{j=1}^{m_2} \varphi_{i_0,j}$, it follows that $G_0 \models \varphi_{i_0,j_0}$.

There are two cases, depending on whether φ_{i_0,j_0} is a monadic NP sentence or a monadic co-NP sentence. If φ_{i_0,j_0} is a monadic NP sentence, then it follows by standard arguments (see, for example, [Fag97]) and our choice of c, r that the spoiler has a winning strategy in the asymmetric (c, r)-game over G_0, G_1. Similarly, if φ_{i_0,j_0} is a monadic co-NP sentence, then the spoiler has a winning strategy in the asymmetric (c, r)-game over G_1, G_0. So in either case the spoiler has a winning strategy in the symmetric (c, r)-game over G_0, G_1, which was to be shown.

We now prove the converse. By standard arguments (such as those in [Fag97]), it is not hard to show that for every structure G_0 and every choice of c, r, there is a sentence $\sigma(G_0, c, r)$ that is a conjunction of sentences of the form $\exists A_1 \cdots \exists A_k \psi$ and of negations of sentences of this form, such that $k = \lceil \log c \rceil$ and ψ is of quantifier depth at most r, and such that the duplicator has a winning strategy in the symmetric (c, r)-game over G_0, G_1 iff $G_1 \models \sigma(G_0, c, r)$. Furthermore, for a given choice of c, r, there are only a finite number of distinct inequivalent such sentences $\sigma(G_0, c, r)$.

Assume now that there are c, r such that whenever $G_0 \in \mathscr{S}$ and $G_1 \in \bar{\mathscr{S}}$, then the spoiler has a winning strategy in the symmetric (c, r)-game over G_0, G_1. Let $\mathscr{S}'$ be a finite subset of $\mathscr{S}$ such that for every $G_0 \in \mathscr{S}$, there is $G_0' \in \mathscr{S}'$ with $\sigma(G_0, c, r)$ equivalent to $\sigma(G_0', c, r)$ (such a set $\mathscr{S}'$ exists, because of our earlier observation that for a given choice of c, r, there are only a finite number of distinct inequivalent sentences of the form $\sigma(G, c, r)$). Let μ be the disjunction $\bigvee_{G_0' \in \mathscr{S}'} \sigma(G_0', c, r)$. To show that $\mathscr{S}$ is in the Boolean closure of monadic NP, we need only show that every member of $\mathscr{S}$ satisfies μ, and no member of $\bar{\mathscr{S}}$ does.

Assume that $G_0 \in \mathscr{S}$. By definition of $\sigma(G_0, c, r)$, it follows easily that $G_0 \models \sigma(G_0, c, r)$. Now $\sigma(G_0, c, r)$ is equivalent to some disjunct $\sigma(G_0', c, r)$ of μ. Hence, $G_0 \models \mu$, as desired. Assume now that $G_1 \in \bar{\mathscr{S}}$. The proof is complete if we show that $G_1 \not\models \mu$. Assume by way of contradiction that $G_1 \models \mu$. Then $G_1 \models \sigma(G_0', c, r)$ for some disjunct $\sigma(G_0', c, r)$ of μ. Therefore, by definition of $\sigma(G_0', c, r)$, the duplicator has a winning strategy in the symmetric (c, r)-game over G_0', G_1. But this contradicts our choice of c, r, since $G_0' \in \mathscr{S}$ and $G_1 \in \bar{\mathscr{S}}$. ▮

We now define a property that we shall show is in the positive first-order closure of monadic NP but not in the Boolean closure of monadic NP. Let G be a directed graph. A subset M of the vertices of G is called a *module* [Spi92] (or an *autonomous set* [MR84] or a *partitive set* [Gol80]) if the members of M are indistinguishable as far as the vertices outside of M are concerned. That is, if u is a vertex of G outside of M, and if v_1, v_2 are in M, then (a) $\langle u, v_1 \rangle$ is an edge of G iff $\langle u, v_2 \rangle$ is an edge of G, and (b) $\langle v_1, u \rangle$ is an edge of G iff $\langle v_2, u \rangle$ is an edge of G. Note that the set of all vertices of G and every singleton set (consisting of a single vertex of G) is automatically a module. An induced subgraph H of G is an *inner factor of G* [Gol80] if the set of vertices of H is a module of G. Intuitively, G is obtained by starting with another graph G' and replacing a fixed vertex x of G' by H. To clarify this intuition, we give some notation (from [Gol80]) that will be useful later. Let H_0 be a graph, with exactly n vertices $v_1, ..., v_n$, and let $H_1, ..., H_n$ be n graphs. We assume for convenience in our definitions that no two graphs H_i, H_j share a vertex; whenever we want to apply the definitions with some of the H_i's being the same graph, we assume that we simply replace the graphs if necessary by isomorphic copies where the universes are disjoint. We write $H_0[H_1, ..., H_n]$ for the result of replacing vertex v_i with the graph H_i, for $1 \leqslant i \leqslant n$, and for each i, j, putting an edge from each vertex of H_i to each vertex of H_j iff $\langle v_i, v_j \rangle$ is an edge of H_0.[4] More formally, if H_i has vertices V_i and edges E_i, for $0 \leqslant i \leqslant n$, then the set of vertices of $H_0[H_1, ..., H_n]$ is $\bigcup_{i=1}^n V_i$, and the set of edges is

$$\left(\bigcup_{i=1}^{n} E_i\right) \cup \left(\bigcup_{i=1}^{n} \bigcup_{j=1}^{n} \{\langle x, y \rangle \mid x \in V_i,\ y \in V_j, \text{ and } \langle v_i, v_j \rangle \in E_0\}\right). \tag{2}$$

[4] In spite of the notation, the graph $H_0[H_1, ..., H_n]$ depends not just on $H_0, H_1, ..., H_n$, but also on the specific ordering $v_1, ..., v_n$ of the vertices of H_0.

We shall also need the definitions of module and inner factor for undirected graphs. The definitions are in the same spirit as the definitions above for directed graphs. A subset M of the vertices of an undirected graph G is a *module* if, for all vertices $u \notin M$ and all vertices $v_1, v_2 \in M$, there is an edge between u and v_1 iff there is an edge between u and v_2. Again, an induced subgraph H of G is an *inner factor* of G if the set of vertices of H is a module of G. For undirected graphs $H_0, H_1, ..., H_n$, the set of edges of $H_0[H_1, ..., H_n]$ is given by (2) where the directed edges $\langle x, y\rangle$ and $\langle v_i, v_j\rangle$ are replaced by the undirected edges (x, y) and (v_i, v_j), respectively.

In general, there are multiple ways to "decompose" a graph into the form $H_0[H_1, ..., H_n]$. In [Gol80], for each such decomposition, the graph H_0 is called the *outer factor*, and the graphs $H_1, ..., H_n$ are called the *inner factors*. Thus, for us, an induced subgraph of a graph is an inner factor precisely if it is an inner factor of some decomposition.

The property we now consider is as follows: "At most one connected component has each of its inner factors connected". It is important to note that this is a property of *undirected graphs*. Therefore, to determine whether a directed graph has this property, we first view it as an undirected graph by ignoring the directions of the edges as explained in Section 2, and then determine if at most one connected component of this undirected graph has each of its inner factors connected. The distinction between directed and undirected graphs is important here, because a set M of vertices of a directed graph G might not be a module of G, but be a module of the undirected graph obtained from G by ignoring the directions of the edges. (As it turns out, however, for the particular directed graphs used in the proof of the next result, the modules of the directed graph are identical to the modules of the associated undirected graph.)

THEOREM 8.3. *The property "At most one connected component has each of its inner factors connected" is not in the Boolean closure of monadic NP.*

Proof. Let $\mathcal{S}$ be the property "At most one connected component has each of its inner factors connected". By Theorem 8.2, we need only show that for every c, r, there are $G_0 \in \mathcal{S}$ and $G_1 \in \bar{\mathcal{S}}$ such that the duplicator has a winning strategy in the symmetric (c, r)-game over G_0, G_1.

So let c, r be given. By Theorem 7.1, there is $a \geqslant 5$ such that if A is a cycle of length a, then the duplicator has a winning strategy in the asymmetric (c, r)-game over $A, A \oplus A$. Again by Theorem 7.1, there is $b \geqslant 5$ such that if B is a cycle of length b, then the duplicator has a winning strategy in the asymmetric (c^a, r)-game over $B, B \oplus B$.

Let A be a cycle of length a, and let B be a cycle of length b. Let S be the graph $B[A, ..., A]$. Thus, intuitively, S is a cycle of cycles. Let D be the graph $B[A \oplus A, ..., A \oplus A]$. Intuitively, D is a cycle of pairs of cycles. We have chosen the names so that S stands for "single" and D for "double".

Let G_0 be the graph consisting of the disjoint union of one copy of S and $rb^{c^{2a}} + 1$ copies of D, and let G_1 be the graph consisting of the disjoint union of two copies of S and $rb^{c^{2a}}$ copies of D. Let $\hat{A}, \hat{D}, \hat{S}$ be the undirected graphs obtained from A, D, S respectively, by ignoring the directions of the edges. It is not hard to see that $G_0 \in \mathcal{S}$ and $G_1 \in \bar{\mathcal{S}}$. This follows from the fairly straightforward fact that the only inner factors of $\hat{S}$ (other than singleton graphs and $\hat{S}$ itself) are copies of $\hat{A}$, and at least one inner factor of $\hat{D}$ is the nonconnected graph $\hat{A} \oplus \hat{A}$.[5]

Since $G_0 \in \mathcal{S}$ and $G_1 \in \bar{\mathcal{S}}$, it follows from Theorem 8.2 that the proof is complete if we show that the duplicator has a winning strategy in the symmetric (c, r)-game over G_0, G_1. As we noted earlier, to prove this, it is sufficient to show that (1) the duplicator has a winning strategy in the asymmetric (c, r)-game over G_0, G_1 and (2) the duplicator has a winning strategy in the asymmetric (c, r)-game over G_1, G_0.

We first show that the duplicator has a winning strategy in the asymmetric (c, r)-game over G_0, G_1. Let G_0' be the subgraph of G_0 consisting of the single copy of S in G_0, and let G_0'' be the subgraph of G_0 consisting of the $rb^{c^{2a}}+1$ copies of D in G_0. Similarly, let G_1' be the subgraph of G_1 consisting of the two copies of S in G_1, and let G_1'' be the subgraph of G_1 consisting of the $rb^{c^{2a}}$ copies of D in G_1. Thus, G_0 is the disjoint union of G_0' and G_0'', and G_1 is the disjoint union of G_1' and G_1''. To show that the duplicator has a winning strategy in the asymmetric (c, r)-game over G_0, G_1, it is clearly sufficient to show that (1) the duplicator has a winning strategy in the asymmetric (c, r)-game over G_0', G_1', and (2) the duplicator has a winning strategy in the asymmetric (c, r)-game over G_0'', G_1''.

We now show that the duplicator has a winning strategy in the asymmetric (c, r)-game over G_0', G_1'. Assume that the spoiler colors G_0' with the c colors. Let us call each coloring of an inner factor A with the c colors a "supercolor". Thus, there are at most c^a supercolors. We think of the spoiler's coloring of G_0' (with c colors) as a supercoloring of B, that is, a coloring of B where each point of B is colored with a supercolor. By definition of b, the duplicator has a winning strategy in the asymmetric (c^a, r)-game over $B, B \oplus B$. The supercoloring that the duplicator would use in his winning strategy to supercolor $B \oplus B$ in the asymmetric (c^a, r)-game over $B, B \oplus B$, based on the spoiler's supercoloring of B, translates into a coloring (using c colors) of G_1'. This is how the duplicator colors G_1'.

In the r-game over the colored G_0', G_1', the duplicator makes use of his winning strategy in the r-game over the supercolored $B, B \oplus B$. Let us refer to the points in B and $B \oplus B$ as *supernodes*. We can think of each supernode as being supercolored with one of the c^a supercolors. By construction, the duplicator has a winning strategy in the r-game over the supercolored $B, B \oplus B$. The duplicator uses this strategy to select supernodes, based on the spoiler's choices of supernodes. "Within" a supernode (that is, on the cycle A that corresponds to that supernode), the duplicator selects corresponding points in the two structures G_0' and G_1'. For example, assume that the spoiler selects a point u in G_0'. This point is on a cycle (call it A_0) that corresponds to some supernode u' of B. Let v' be the supernode that the duplicator would now pick in $B \oplus B$ in his winning r-game strategy over the supercolored $B, B \oplus B$, based on the spoiler selecting u'. The duplicator then selects a point v in G_1' that lies on the cycle (call it A_1) that corresponds to the supernode v', and in particular selects that point v that is in the same position in

[5] Here we use the fact that $a \geqslant 5$ and $b \geqslant 5$, since, for example, if $b = 4$, then the graph consisting of two copies of $\hat{A}$ that are "nonadjacent" in $\hat{S}$ together form an inner factor of $\hat{S}$.

A_1 as u is in A_0. It is straightforward to see that the duplicator thereby wins the r-game over the colored G'_0, G'_1. This completes the proof that the duplicator has a winning strategy in the asymmetric (c, r)-game over G'_0, G'_1.

We now show that the duplicator has a winning strategy in the asymmetric (c, r)-game over G''_0, G''_1. Assume that the spoiler colors G''_0 with the c colors. Since there are at most c^{2a} ways to color $A \oplus A$ (with the c colors), there are at most $b^{c^{2a}}$ ways to color D. Since there are $rb^{c^{2a}}+1$ copies of D in G''_0, it follows easily that there exists a set Y of $r+1$ copies of D in G''_0 that are all colored the same. Let Z be the set of all but one of the $rb^{c^{2a}}+1$ copies of D in G''_0, where the missing copy of D is one of the members of Y. The duplicator then colors the $rb^{c^{2a}}$ copies of D in G''_1 by mimicking the colorings of the $rb^{c^{2a}}$ members of Z.

In the r-game over the colored G''_0, G''_1, the duplicator essentially mimics the moves of the spoiler. The point is that since the set Y as defined above contains at least $r+1$ members, there is some member of Y with no point selected from it in the r rounds. The simple details are left to the reader. This concludes the proof that the duplicator has a winning strategy in the asymmetric (c, r)-game over G''_0, G''_1, and hence concludes the proof that the duplicator has a winning strategy in the asymmetric (c, r)-game over G_0, G_1.

We now show that the duplicator has a winning strategy in the asymmetric (c, r)-game over G_1, G_0. We "divide up" G_0 into subgraphs G'_0 and G''_0 in a different manner than before. Specifically, let G'_0 be the subgraph of G_0 consisting of a single copy of D in G_0. Let G''_0 be the subgraph of G_0 consisting of the rest, that is, the single copy of S and $rb^{c^{2a}}$ copies of D in G_0. We also divide up G_1 into subgraphs G'_1 and G''_1 in a different manner than before. Specifically, let G'_1 be the subgraph of G_1 consisting of a single copy of S in G_1. Let G''_1 be the subgraph of G_1 consisting of the rest, that is, one copy of S and $rb^{c^{2a}}$ copies of D in G_1. As before, G_0 is the disjoint union of G'_0 and G''_0, and G_1 is the disjoint union of G'_1 and G''_1.

Like before, to show that the duplicator has a winning strategy in the asymmetric (c, r)-game over G_1, G_0, it is clearly sufficient to show that (1) the duplicator has a winning strategy in the asymmetric (c, r)-game over G'_1, G'_0, and (2) the duplicator has a winning strategy in the asymmetric (c, r)-game over G''_1, G''_0. The proof of (2) is trivial, since G''_1 and G''_0 are isomorphic. So we need only prove (1). That is, we need only prove that the duplicator has a winning strategy in the asymmetric (c, r)-game over S, D.

Now intuitively, S is a cycle (of length b) of cycles A, whereas D is a cycle (of the same length b) of pairs $A \oplus A$ of cycles. It is easy to see how the duplicator can use the winning strategy in the asymmetric (c, r)-game over $A, A \oplus A$ to get a winning strategy in the asymmetric (c, r)-game over S, D. Essentially, the duplicator views the (c, r)-game over S, D as b disjoint copies of the (c, r)-game over $A, A \oplus A$ played in parallel. The straightforward details are left to the reader. ∎

We now show that the property "At most one connected component has each of its inner factors connected" belongs to the positive first-order closure of monadic NP, and, moreover, can be expressed by a sentence of the form $\forall x \, \forall y \psi$ where ψ is a monadic NP formula.

Recall that, for a given graph G, the notation $ind(X)$ stands for the subgraph of G induced by the set X of points, and $CC(x)$ stands for the connected component of G that contains the point x.

THEOREM 8.4. *The property "At most one connected component has each of its inner factors connected" can be expressed by a sentence of the form* $\forall x \forall y \psi$ *where* ψ *is a monadic NP formula.*

Proof. Clearly, the following high-level sentence expresses the property stated in the theorem:

$\forall x \forall y$("x and y are path-connected"

$\vee$ "$CC(x)$ contains a nonconnected inner factor"

$\vee$ "$CC(y)$ contains a nonconnected inner factor").

It suffices to show that each of the three terms in the disjunction can be expressed by a monadic NP formula. We have already shown in Section 4 that the first term can be so expressed, so we only have to describe a monadic NP formula for the second term (the same formula works for the third term by replacing x with y).

Again at a high level, "$CC(x)$ contains a nonconnected inner factor" is expressed by the following:

There exists a set M of points such that:

(1) M is a module;

(2) $ind(M)$ is nonconnected;

(3) x is path-connected to some point in M; and

(4) for every two distinct points v_1 and v_2 in M, the points v_1 and v_2 are connected by a path of length two.

We first argue that this expresses the desired property. Say first that $CC(x)$ contains a nonconnected inner factor. Let M be the set of points of some nonconnected inner factor contained in $CC(x)$. Clearly, (1), (2), and (3) are satisfied. To see that (4) is satisfied, note that M is properly contained in the set V_x of points of $CC(x)$, because $ind(M)$ is nonconnected whereas $CC(x)$ is connected. Let v_1 and v_2 be distinct points of M. Since M is properly contained in V_x and $CC(x)$ is connected, there must be a point $v_3 \in M$ and a point $u \in V_x - M$ with an edge between v_3 and u. The definition of module implies that there is an edge between u and v_1 and an edge between u and v_2, so v_1 and v_2 are connected by a path of length two.

Say now that (1)–(4) hold for a set M. It is easy to see that (3) and (4) imply that $ind(M)$ is contained in $CC(x)$. So, by (1) and (2), $ind(M)$ is a nonconnected inner factor of $CC(x)$.

To complete the proof, we note that each of (1)–(4) can be expressed by a monadic NP formula. For (2), "$ind(M)$ is nonconnected", we have shown this in Section 4. Obviously, (1) and (4) can be expressed in first-order. A monadic NP formula expressing (3), "x is path-connected to some point in M", is similar to the formula described in Section 4 for "x and y are path-connected". Thus, (3) is expressed by $\exists A \, \exists w(w \in M \wedge path(A, x, w))$, where $path(A, x, w)$ is the first-order formula described in Section 4, that says that A is a strict path between x and w. ∎

The following theorem is an immediate consequence of Theorems 8.3 and 8.4.

THEOREM 8.5. *The property "At most one connected component has each of its inner factors connected" is in the positive first-order closure of monadic NP but not in the Boolean closure of monadic NP.*

This theorem immediately implies Theorem 1.1, which is one of our main separation results.

9. THE POWER OF CLOSED MONADIC NP

Properties in closed monadic NP are defined by sentences of the form $\mathbf{Q}\varphi$, where the prefix $\mathbf{Q}$ can have an arbitrary interleaving of first-order quantifiers and existential unary second-order quantifiers, and where φ is first-order. Do we gain expressive power by allowing interleaving? In particular, is there a property in closed monadic NP that is not in the positive first-order closure of monadic NP? In this section, we show that the answer is yes: we indeed gain expressive power by allowing interleaving. In fact, we actually prove a stronger result. We show that there is a property in closed monadic NP that is not in the first-order closure of monadic NP.

Let $\mathscr{P}_1$ be the property "At most one connected component has each of its inner factors connected". In Section 8, we showed that $\mathscr{P}_1$ is not in the Boolean closure of monadic NP, but that it is in the positive first-order closure of monadic NP. We now define a new undirected graph property $\mathscr{P}_2$ to be "There is a cycle C of length $n \geqslant 5$ and graphs $H_1, ..., H_n$ such that the number of H_i's that satisfy $\mathscr{P}_1$ is even, and the graph is $C[H_1, ..., H_n]$". Note that this property is well-defined since it does not matter how the graphs $H_1, ..., H_n$ are substituted for the points of C; all that matters is that the number of H_i's that satisfy $\mathscr{P}_1$ is even. In this section, we show that property $\mathscr{P}_2$ is in closed monadic NP, but not in the first-order closure of monadic NP.

Before we prove that property $\mathscr{P}_2$ is not in the first-order closure of monadic NP, we need some preliminary results. First, we give a game that characterizes the first-order closure of monadic NP. In this game, there are first-order rounds, followed by a coloring round, followed by first-order rounds. Let G_0, G_1 be structures, and let r_1, c, r_2 be positive integers (where r_1 represents the number of initial first-order rounds, c the number of colors, and r_2 the number of remaining first-order rounds). We call this game the *symmetric* (r_1, c, r_2)*-game over* G_0, G_1. The rules are as follows.

1. There are r_1 rounds where the spoiler selects a point in G_0 or G_1 and the duplicator selects a point in the other structure.
2. The spoiler colors either G_0 or G_1 with the c colors.
3. The duplicator colors the other structure with the c colors.

4. There are r_2 more rounds where the spoiler selects a point in G_0 or G_1 and the duplicator selects a point in the other structure.

The winner is decided as before. Again, the isomorphism must respect colors.

The next proposition will be used to characterize the first-order closure of monadic NP in terms of the symmetric (r_1, c, r_2)-game.

PROPOSITION 9.1. *Every FO(MNP) formula is equivalent to a PFO(BOOL (MNP)) formula, that is, a formula of the form* $\mathbf{P}\psi$ *where* $\mathbf{P}$ *consists of first-order quantifiers and* ψ *is a Boolean combination of monadic NP formulas.*

Proof. It is enough to show that (i) the PFO(BOOL(MNP)) formulas are closed under first-order quantification, and (ii) if φ_1 and φ_2 are PFO(BOOL (MNP)) formulas, then $\neg\varphi_1$ and $\varphi_1 \wedge \varphi_2$ are each equivalent to a PFO(BOOL (MNP)) formula. Since (i) is obvious, we complete the proof by showing (ii). For $i = 1, 2$, write $\varphi_i = \mathbf{P}_i\psi_i$, where $\mathbf{P}_i$ consists of first-order quantifiers and ψ_i is a BOOL(MNP) formula. Now $\neg\varphi_1$ is equivalent to the PFO(BOOL(MNP)) formula $\mathbf{P}'_1\neg\psi_1$, where $\mathbf{P}'_1$ is obtained from $\mathbf{P}_1$ by changing each existential quantifier to a universal quantifier, and vice versa. And $\varphi_1 \wedge \varphi_2$ is equivalent to the PFO(BOOL(MNP)) formula $\mathbf{P}'_1\mathbf{P}'_2(\psi'_1 \wedge \psi'_2)$, where $\mathbf{P}'_1$, ψ'_1, $\mathbf{P}'_2$, and ψ'_2 are obtained from $\mathbf{P}_1$, ψ_1, $\mathbf{P}_2$, and ψ_2, respectively, by renaming variables in such a way that the variables in $\mathbf{P}'_1$ do not appear in ψ'_2, and the variables in $\mathbf{P}'_2$ do not appear in ψ'_1. ∎

We can now characterize the first-order closure of monadic NP in terms of the symmetric (r_1, c, r_2)-game.

THEOREM 9.2. $\mathscr{S}$ *is in the first-order closure of monadic NP iff there are* r_1, c, r_2 *such that whenever* $G_0 \in \mathscr{S}$ *and* $G_1 \in \bar{\mathscr{S}}$, *then the spoiler has a winning strategy in the symmetric* (r_1, c, r_2)*-game over* G_0, G_1.

Proof. Given Proposition 9.1, the proof is similar to that of Theorem 8.2 by using fairly standard modifications, which are omitted. Roughly, the idea is that, given a PFO(BOOL(MNP)) sentence $\mathbf{P}\psi$ as above, the first r_1 point-selecting rounds correspond to the first-order quantifiers in $\mathbf{P}$, and the symmetric coloring round and the final r_2 point-selecting rounds correspond to the BOOL(MNP) formula ψ as in the proof of Theorem 8.2. ∎

We now develop machinery that we will need to prove that property $\mathscr{P}_2$ is not in the first-order closure of monadic NP. We begin with a lemma that we will make use of several times. A *chain* is a (directed) graph with points $\alpha_0, ..., \alpha_{n-1}$ such that the edges are precisely $\langle\alpha_0, \alpha_1\rangle, \langle\alpha_1, \alpha_2\rangle, ..., \langle\alpha_{n-2}, \alpha_{n-1}\rangle$. The *length* of the chain is n. Note that this length n is equal to the number of points, not the number of edges. We say that α_i *precedes* α_j on the chain if $i < j$. The *left endpoint* of the chain is α_0, the *next-to-left endpoint* is α_1, the *right endpoint* is α_{n-1}, and the *next-to-right endpoint* is α_{n-2}. Let A and B be chains. In an r-game over A, B, let a_i (resp., b_i) be the point selected during round i in A (resp., B), for $1 \leqslant i \leqslant r$. A duplicator strategy is said to be *order-respecting* if a_i precedes a_j in A iff b_i precedes b_j in B, for each i, j with $1 \leqslant i \leqslant r$ and $1 \leqslant j \leqslant r$. A duplicator strategy is said to be *endpoint-respecting* if a_i is the left (resp., next-to-left, right, next-to-right) endpoint of A iff

b_i is the left (resp., next-to-left, right, next-to-right) endpoint of B, for each i with $1 \leqslant i \leqslant r$.

LEMMA 9.3. *Let A and B be chains each of length greater than 2^{r+1}. Then the duplicator has a winning strategy in the r-game over A, B such that the strategy is both order-respecting and endpoint-respecting.*

Proof. Let A and B be chains each of length greater than 2^{r+1}. Define numbers $t_r, ..., t_0$ by $t_r = 1$ and $t_i = 2t_{i+1} + 1$, for $i = r-1, ..., 0$. It is easy to show by reverse induction on i that $t_i = 2^{r-i+1} - 1$ for all i. Let a_{-1} (resp., b_{-1}) denote the left endpoint of A (resp., B), and let a_0 (resp., b_0) denote the right endpoint of A (resp., B). We describe a strategy for the duplicator such that, for $i = 0, 1, ..., r$, after i rounds have been played the following invariant holds. For each j, k with $-1 \leqslant j \leqslant i$ and $-1 \leqslant k \leqslant i$:

1. a_j precedes a_k in A iff b_j precedes b_k in B (in particular, $a_j = a_k$ iff $b_j = b_k$); and

2. either the distance between a_j and a_k is the same as the distance between b_j and b_k, or else both distances are greater than t_i.

Note that the invariant holds for $i = 0$ because $t_0 < 2^{r+1}$ and because of our assumption on the lengths of A and B. Note that this is an endpoint-respecting winning strategy because $t_r = 1$. It is immediate from the first part of the invariant that the strategy is order-respecting.

For some i with $0 \leqslant i < r$, say that i rounds have been played, and that the invariant holds. Say that the spoiler selects a point a_{i+1} from A (the case where the spoiler selects from B is completely symmetric). If $a_{i+1} = a_j$ for some $j \leqslant i$, then the duplicator selects $b_{i+1} = b_j$. Say then that a_{i+1} lies strictly between adjacent points a_j and a_k where $j, k \leqslant i$ (when we say that a_j and a_k are *adjacent*, here we mean that there is no a_ℓ with $1 \leqslant \ell \leqslant i$ that is strictly between a_j and a_k). Let $\delta(x, y)$ denote the distance between points x, y. If $\delta(a_j, a_k) = \delta(b_j, b_k)$, then the duplicator selects b_{i+1} between b_j and b_k such that $\delta(b_j, b_{i+1}) = \delta(a_j, a_{i+1})$. If $\delta(a_j, a_k) \neq \delta(b_j, b_k)$, then the distances $\delta(a_j, a_k)$ and $\delta(b_j, b_k)$ are both greater than $t_i = 2t_{i+1} + 1$. If a_{i+1} is within distance t_{i+1} from a_j (resp., a_k), then the duplicator selects b_{i+1} between b_j and b_k to be at the same distance from b_j (resp., b_k). The final case is that a_{i+1} is farther than t_{i+1} from both a_j and a_k. Since $\delta(b_j, b_k) > t_i = 2t_{i+1} + 1$, there is a point between b_j and b_k that is farther than t_{i+1} from both b_j and b_k. Let b_{i+1} be such a point; this is the duplicator's selection. In each case, the invariant holds for $i + 1$. ∎

As an intermediate step in showing that property $\mathscr{P}_2$ is not in the first-order closure of monadic NP, we now give a lemma about a certain type of game played on two-colored cycles. A *black/white cycle* is a (directed) cycle with each point colored either black or white. We think of the direction of the edges as giving a "clockwise" orientation to a directed cycle. Let r, t, M be positive integers, and let A and B be black/white cycles. The rules of the *(r, t, M)-black/white game over A, B* are as follows. The game has r rounds. In each round the spoiler selects a point in one structure and the duplicator responds by selecting a point in the other structure. Let a_i (resp., b_i) be the point selected during round i in A (resp., B), for $1 \leqslant i \leqslant r$. The duplicator wins if:

1. a_i has the same color as b_i, for $1 \leqslant i \leqslant r$;
2. $a_i = a_j$ iff $b_i = b_j$, for all i, j;
3. The sequences $a_1, ..., a_r$ and $b_1, ..., b_r$ appear in the same order around the two cycles. More precisely, there is a permutation $\pi: \{1, ..., r\} \to \{1, ..., r\}$ with $\pi(1) = 1$ such that, starting at a_1 (resp., b_1) and moving in the clockwise direction around the cycle A (resp., B), the points occur in the order $a_{\pi(1)}, ..., a_{\pi(r)}$ (resp., $b_{\pi(1)}, ..., b_{\pi(r)}$)[6]; and
4. For each pair of adjacent (in the π order) selected points in one structure and the corresponding pair in the other structure, either the distance between them is the same in both structures, or both distances are "large" (greater than t) and congruent to each other mod M. More precisely, for $1 \leqslant i \leqslant r$, let ℓ_i (resp., ℓ'_i) be the distance from $a_{\pi(i-1)}$ to $a_{\pi(i)}$ (resp., from $b_{\pi(i-1)}$ to $b_{\pi(i)}$) moving in the clockwise direction, where $\pi(0) = \pi(r)$. Then, for $1 \leqslant i \leqslant r$, either

   (a) $\ell_i = \ell'_i$, or

   (b) $\ell_i > t$, and $\ell'_i > t$, and $\ell_i \equiv \ell'_i \bmod M$.

LEMMA 9.4. *Let r, t, M be positive integers. There are arbitrarily large black/white cycles A and B of the same size such that an even (resp., odd) number of points of A (resp., B) are colored white, and the duplicator has a winning strategy in the (r, t, M)-black/white game over A, B.*

Proof. Let $N = tM$. In particular, $N \geqslant t$ and $N \equiv 0 \bmod M$. By increasing t if necessary, we can make N arbitrarily large. Define a *colored chain* to be a chain where each of the points is colored black or white. Define C_0 to be a colored chain of N points, where every point is colored black. Define C_1 to be a colored chain of N points, where the first point (α_0 above) is colored white, and the remaining points are colored black. Thus, C_i contains exactly i white points, for $i = 0, 1$. Let k be an odd number (whose existence is guaranteed by Lemma 9.3) such that if A and B are chains of length at least k, then the duplicator has an order-respecting, endpoint-respecting winning strategy in the r-game over A, B. Let A be the black/white cycle constructed by "stringing together" $k+1$ consecutive C_1's, followed by k consecutive C_0's. These $2k+1$ chains are strung together by placing an edge from the right endpoint of each chain to the left endpoint of the next chain, and finally by placing an edge from the right endpoint of the last copy of C_0 to the left endpoint of the first copy of C_1. Let B be defined similarly, except that there are k consecutive C_1's, followed by $k+1$ consecutive C_0's. Since k is odd, it follows that A has an even number of white points, and B has an odd number of white points. We now show that the duplicator has a winning strategy in the (r, t, M)-black/white game over A, B.

Similarly to before, certain groups of points of A and B are viewed as *supernodes*. We think of A as a cycle consisting of a chain of $k+1$ supernodes that correspond

[6] We note that some of the a_i's (and similarly the b_i's) might be identical. In the definition of π, the "order" of identical points can be chosen arbitrarily.

to C_1 (let us call this chain A'), followed by a chain of k supernodes that correspond to C_0 (let us call this chain A''). Similarly, we think of B as a cycle consisting of a chain of k supernodes that correspond to C_1 (let us call this chain B'), followed by a chain of $k+1$ supernodes that correspond to C_0 (let us call this chain B''). By our choice of k, the duplicator has an order-respecting, endpoint-respecting winning strategy in the r-game over A', B' (where we think of the spoiler and duplicator as selecting supernodes in each round), and similarly the duplicator has an order-respecting, endpoint-respecting winning strategy in the r-game over A'', B''. Note from our definition of "endpoint-respecting" that, for example, if the spoiler selects, say, the left endpoint of A' (that is, the leftmost supernode), then the duplicator must select the left endpoint of B'. These winning strategies can be combined into a winning strategy in the r-game over A, B, where again we think of the players as selecting supernodes. Finally, this winning strategy can be converted into a winning strategy in the "real" r-game over A, B (where points, not supernodes, are selected), by having the duplicator always select corresponding points within supernodes. Thus, for example, if the spoiler selects the jth point from the left in the chain (C_0 or C_1) corresponding to the supernode a of A, and if the winning strategy on supernodes would cause the duplicator to select the supernode b of B in response to the spoiler's selection of the supernode a of A, then the duplicator selects the jth point from the left in the chain corresponding to the supernode b of B. To show that this gives a winning strategy for the duplicator, note first that the first three conditions for the duplicator to win the black/white game hold. We now show that the fourth condition also holds. Let us adopt the notation of the fourth condition. Since the duplicator always selects corresponding points within supernodes, and since the length $N = Mt$ of each of the chains C_0, C_1 associated with a supernode is congruent to 0 mod M, it follows that $\ell_i \equiv \ell'_i \bmod M$. There are now three cases.

Case 1. $a_{\pi(i-1)}$ and $a_{\pi(i)}$ are in the same supernode (that is, are both in a chain corresponding to the same supernode). Then also $b_{\pi(i-1)}$ and $b_{\pi(i)}$ are in the same supernode, so from the duplicator's strategy, we have $\ell_i = \ell'_i$.

Case 2. $a_{\pi(i-1)}$ and $a_{\pi(i)}$ are in adjacent supernodes. Then from the duplicator's strategy, $b_{\pi(i-1)}$ and $b_{\pi(i)}$ are also in adjacent supernodes, and from the duplicator's strategy, again $\ell_i = \ell'_i$.

Case 3. $a_{\pi(i-1)}$ and $a_{\pi(i)}$ are in different supernodes that are not adjacent. Then from the duplicator's strategy, $b_{\pi(i-1)}$ and $b_{\pi(i)}$ are in different supernodes that are not adjacent. So, since $N \geqslant t$, it follows that $\ell_i > t$ and $\ell'_i > t$. As we already showed, $\ell_i \equiv \ell'_i \bmod M$. This completes the proof that the fourth condition holds for the duplicator to win. So indeed, the duplicator has a winning strategy. ∎

We need one more lemma.

LEMMA 9.5. *Let c, r be positive integers. There are positive integers t, M such that whenever A, B are chains each of length greater than t and whose lengths are congruent mod M, then the duplicator has an endpoint-respecting winning strategy in the symmetric (c, r)-game over A, B.*

Proof. We first dispose of the issue of requiring the strategy to be endpoint-respecting. Consider a symmetric $(c, r+2)$-game over A, B. It is easy to see that any winning strategy for the duplicator in this game must be endpoint-respecting at least through the first r point-selecting rounds. Suppose, for example, that after i rounds, where $i \leqslant r+1$, the point a_i is the left endpoint α_0 of A and b_i is some point β_k other than the left endpoint of B. Then on the next round the spoiler can select the point β_{k-1} such that there is an edge $\langle \beta_{k-1}, \beta_k \rangle$, and the duplicator cannot respond in a winning strategy since there is no edge directed into α_0. Thus, to prove the lemma it suffices to replace r with $r+2$ and show that the duplicator has some winning strategy in the symmetric $(c, r+2)$ game over A, B. For simplicity in the following, we assume that r has already been replaced by $r+2$ and do not mention it further.

Let $f=2$ (the maximal degree of a point in a chain), and let d, m be as in Theorem 6.1. We assume without loss of generality that the number c of colors is at least 2. Assume that the spoiler colors A with the c colors. (The case where the spoiler colors B is completely symmetric.) We shall now mimic part of the proof of Lemma 7.2.

Let $\alpha_0, \alpha_1, ..., \alpha_{n-1}$ be the points in order in A, so there is an edge from α_i to α_{i+1} for $0 \leqslant i < n-1$. Let $\chi(\alpha_i)$ denote the color of α_i, and let $\chi(A)$ denote the colored version of A. When $n \geqslant 2d$ then the d-type of the point α_i is described by a vector of up to $2d-1$ colors[7] as in (1) in Section 7, along with a description of whether the point is less than distance d from one of the endpoints of the chain, and if so, which endpoint and what the distance is. Let τ_i denote the d-type of α_i.

Let N be the number of d-types. Let $M \geqslant m$ be chosen as z was in the proof of Lemma 7.2, that is, so that M is divisible by every integer from 2 to dN. Let $\Delta = MN+1$, let $t_0 = (m+\Delta) N\Delta$, and let $t = t_0 + \Delta$. As before, say that d-type τ is *rare* if it occurs less than m times as a d-type of a point of $\chi(A)$. Say that τ is *frequent* if it is not rare.

There are two cases, depending on whether or not the length of the chain A is less than the length of the chain B.

Case 1. *The length of A is less than the length of B.* By assumption, the length of A is greater than t. We shall use only the fact that the length of A is greater than t_0, rather than the stronger fact that it is greater than t; the fact that we are making this weaker assumption will be important later on. It is straightforward to see that we have selected t_0 sufficiently large that the argument in the proof of Lemma 7.2 can be used to show that there must be two points α_p and α_q with $0 \leqslant p < q < n$ such that: (1) $\tau_p = \tau_q$; (2) τ_i is frequent for all i with $p \leqslant i \leqslant q$; and (3) $d \leqslant q-p \leqslant dN$. Let I be the colored chain corresponding to the interval $[\alpha_p, \alpha_{q-1}]$. The duplicator should color B by taking the coloring of A and "replacing" I by multiple repeated copies of I. More precisely, the duplicator should color B so that the colors, in order, are

$$\chi(\alpha_0), ..., \chi(\alpha_{p-1}), (\chi(\alpha_p), ..., \chi(\alpha_{q-1}))^h, \chi(\alpha_q), \chi(\alpha_{q+1}), ..., \chi(\alpha_{n-1}),$$

[7] We say "up to", since if the point is near an endpoint of the chain, there are less points in the neighborhood of radius d about the point.

where $(\chi(\alpha_p), ..., \chi(\alpha_{q-1}))^h$ represents the coloring of the interval $[\alpha_p, \alpha_{q-1}]$ repeated h times, and where h is chosen so that the total length of the whole coloring is equal to the length of B. (This choice of h is possible because the lengths of A and B are congruent mod M, and because M is divisible by $q-p$.) As in the proof of Lemma 7.2, the colored versions of A and B are (d, m)-equivalent, because the additional d-types produced by repeating I are all frequent. Therefore, by Hanf's technique (Theorem 6.1), the duplicator can now win the remaining first-order game.

Case 2. *The length of A is at least the length of B.* Recall that $\chi(A)$ is the colored version of A. Let D be a colored chain (colored with the c colors) such that

(1) $D \sim_r \chi(A)$;

(2) The length of D is at least t_0;

(3) the length of D and the length of $\chi(A)$ are congruent mod M; and

(4) the length of D is as small as possible, subject to (1), (2), and (3).

This definition makes sense, because there is some D that satisfies (1), (2), and (3), namely $\chi(A)$.

We now show that the length of D is less than $t = t_0 + \Delta$. Assume not; we shall derive a contradiction. Let $\delta_0, \delta_1, ..., \delta_{v-1}$ be the points in order in D, so there is an edge from δ_i to δ_{i+1} for $0 \leqslant i < v-1$. Let us now say that d-type τ is *rare'* if it occurs less than $m + \Delta$ times as a d-type of a point of D. The differences between rare and rare' are: (i) rare refers to d-types in $\chi(A)$ whereas rare' refers to d-types in D, and (ii) for rare', we use $m + \Delta$ in the definition rather than m. Say that τ is *frequent'* if it is not rare'. We shall refer to a point in D as rare' (resp., frequent') if its d-type is rare' (resp., frequent').

Let us say that a pair of rare' points is *consecutive* if there are no rare' points between them. We now show that between every consecutive pair of rare' points there are less than Δ frequent' points. Assume not; then there is an interval of length Δ in D consisting only of frequent' points. Since $\Delta = MN + 1$, and since N is the number of d-types, there is some d-type that occurs at least $M + 1$ times in this interval. Let $\delta_{i_0}, ..., \delta_{i_M}$ be $M + 1$ distinct points in the interval with the same d-type, with $i_0 < i_1 < \cdots < i_M$. We now show that there are two numbers in the set $\{i_0, i_1, ..., i_M\}$ that are congruent mod M. Consider the differences between i_0 and each of $i_1, i_2, ..., i_M$, that is, $i_1 - i_0, i_2 - i_0, ..., i_M - i_0$. If these differences are all distinct mod M, then one of these differences, say $i_j - i_0$, is congruent to 0 mod M, so $i_j \equiv i_0$ mod M, as desired. So assume that two of these differences are equal mod M, say $i_j - i_0 \equiv i_k - i_0$ mod M. But then $i_j \equiv i_k$ mod M, as desired. Thus, in either case, there are j, k with $j < k$ such that $i_j \equiv i_k$ mod M. Define the colored chain E by deleting the interval $[\delta_{i_j}, \delta_{i_k})$ from D. Thus, E contains the points

$$\delta_0, ..., \delta_{i_j - 1}, \delta_{i_k}, \delta_{i_k + 1}, ..., \delta_{v-1},$$

where each point of E inherits its color from D. By the definition of frequent', the d-type of every point in the deleted interval occurs at least $m + \Delta$ times in D. Since the length of the deleted interval is at most Δ, each such d-type still occurs at least

m times in E. It follows that D and E are (d, m)-equivalent, so by Theorem 6.1, it follows that $E \sim_r \chi(A)$. Since $i_j \equiv i_k \bmod M$, it follows that the length of E is congruent to the length of $D \bmod M$, and hence congruent to the length of $\chi(A) \bmod M$. The length of E is at least equal to the length of D minus Δ. Since by assumption the length of D is at least $t_0 + \Delta$, it follows that the length of E is at least t_0. But this violates minimality of D. This contradiction shows that between every consecutive pair of rare′ points in D there are less than Δ frequent′ points.

Since N is the number of d-types, there are at most $(m + \Delta - 1)\, N$ rare′ points in D (which clearly include the endpoints of D). We just showed that between every consecutive pair of rare′ points in D there are less than Δ frequent′ points. So the total number of points in D is at most $(m + \Delta - 1)\, N\Delta + 1$, which is less than t_0. This contradicts our assumption that the length of D is at least $t_0 + \Delta$. So the length of D is less than $t_0 + \Delta$.

We just showed that the length of D is less than $t = t_0 + \Delta$. In particular, the length of D is less than the length of B. We are now in the situation of the proof of Case 1, where the length of D is at least t_0, the length of D is less than the length of B, and the lengths of B and D are congruent mod M. By the argument in Case 1, the duplicator can color B (to obtain $\chi(B)$) so that $\chi(B) \sim_r D$. Since also $D \sim_r \chi(A)$, it follows that $\chi(A) \sim_r \chi(B)$. Thus, the duplicator can win the remaining r-round game. ∎

We can now prove the key lower bound of this section.

THEOREM 9.6. *Property $\mathscr{P}_2$ is not in the first-order closure of monadic NP.*

Proof. By Theorem 9.2, we need only show that for each choice of r_1, c, r_2, there are H_0, H_1 such that H_0 has property $\mathscr{P}_2$ and H_1 does not, and the duplicator has a winning strategy in the symmetric (r_1, c, r_2)-game over H_0, H_1. By Theorems 8.2 and 8.3, there are G_0, G_1 such that G_0 has property $\mathscr{P}_1$ and G_1 does not, and such that the duplicator has a winning strategy in the symmetric (c, r_2)-game over G_0, G_1. Let us call each coloring of G_0 or G_1 a "supercolor". Let c' be the total number of supercolors. Find t, M from Lemma 9.5 where the roles of c, r in the statement of Lemma 9.5 are played by c', r_2. Let A, B be black/white cycles of length $n \geqslant 5$ guaranteed by Lemma 9.4 where the roles of r, t, M in the statement of Lemma 9.4 are played by r_1, t, M. Let $\alpha_0, \alpha_1, ..., \alpha_{n-1}$ be the points in order around A. Each point is colored either white or black. Let A' be the graph obtained from A by ignoring the colors. Let H_0 be the graph $A'[A_0, ..., A_{n-1}]$, where A_i is G_0 if α_i is white in A, and A_i is G_1 if α_i is black in A. Intuitively, each white point in A is replaced by G_0, and each black point in A is replaced by G_1. Similarly, let $\beta_0, \beta_1, ..., \beta_{n-1}$ be the points in order around B, and let B' be the graph obtained from B by ignoring the colors. Let H_1 be the graph $B'[B_0, ..., B_{n-1}]$, where B_j is G_0 if β_j is white in B, and B_j is G_1 if β_j is black in B. Thus, H_0 has property $\mathscr{P}_2$, but H_1 does not. The proof is complete if we show that the duplicator has a winning strategy in the symmetric (r_1, c, r_2)-game over H_0, H_1.

In the first r_1 rounds the duplicator takes advantage of his winning strategy in the (r_1, t, M)-black/white game over A, B, as guaranteed by Lemma 9.4. Let us refer to the points in A and B as *supernodes*. The duplicator uses his strategy from

Lemma 9.4 to select supernodes, based on the spoiler's choices of supernodes. For example, when the spoiler selects a point in H_0 that is in the graph A_i corresponding to the supernode α_i, let us for now think of this as if the spoiler selected the supernode α_i. The duplicator then selects the supernode β_j corresponding to his strategy as guaranteed by Lemma 9.4. It is important to note that the graph B_j corresponding to the supernode β_j in H_1 is then the same as the graph A_i corresponding to the supernode α_i in H_0: this is because β_j is white iff α_i is white. "Within" a supernode (for example, on the graph B_j that corresponds to the supernode β_j of B), the duplicator then selects the same point as the point within A_i that the spoiler selected in H_0. This describes the duplicator's strategy in the first r_1 rounds.

We now describe the duplicator's coloring strategy. It is convenient to think for now as if the spoiler and duplicator selected (certain) supernodes in the first r_1 rounds, and are now preparing to supercolor (all) the supernodes.

As in the definition of the (r, t, M)-black/white game, let $a_{\pi(1)}, ..., a_{\pi(r_1)}$ (resp., $b_{\pi(1)}, ..., b_{\pi(r_1)}$) be the selected supernodes in clockwise order around A (resp., B). In the coloring round, let us say that the spoiler colors H_0 (a symmetric argument would be used if the spoiler colored H_1). We think of this as a supercoloring of A. The idea now is that for each i with $1 \leqslant i \leqslant r_1$, the duplicator supercolors B within the clockwise interval $[b_{\pi(i)}, b_{\pi(i+1)})$ based on the spoiler's coloring of the clockwise interval $[a_{\pi(i)}, a_{\pi(i+1)})$, using the endpoint-respecting strategy guaranteed by Lemma 9.5 (of course, the indices i are mod r_1). The reason for requiring the strategy to be endpoint-respecting is so that the strategies used for different intervals are consistent at the places where two intervals meet. In particular, since we have assumed $r \geqslant 1$ in Lemma 9.5, $a_{\pi(i)}$ is supercolored the same as $b_{\pi(i)}$ for all i. We now describe how the spoiler actually colors the points of H_1 based on the supercoloring of B. Let β_j be a point of B, and assume that B_j is G_0 (so that β_j is white in B); the argument is symmetric if B_j is G_1 (so that β_j is black in B). What do we mean when we say that the duplicator supercolors β_j with the supercolor γ? Recall that each supercolor represents a coloring of G_0 or G_1. There are two cases, depending on whether γ represents a coloring of G_0 or a coloring of G_1. If γ represents a coloring of G_0, then the duplicator simply colors the inner factor B_j (a copy of G_0) at β_j using γ. So assume that γ represents a coloring of G_1. Now G_0, G_1 were selected so that the duplicator has a winning strategy in the symmetric (c, r_2)-game over G_0, G_1. The duplicator colors the inner factor B_j at β_j as he would color G_0 in his winning strategy in the symmetric (c, r_2)-game over G_0, G_1 if the spoiler were to color G_1 with γ.

We now describe the duplicator's strategy in the final r_2 rounds. Again, we first think in terms of supernodes. Assume, for example, that the spoiler selects a supernode in the clockwise interval $[a_{\pi(i)}, a_{\pi(i+1)})$; then the duplicator selects a supernode in the corresponding clockwise interval $[b_{\pi(i)}, b_{\pi(i+1)})$, where his selection is based on his endpoint-respecting winning strategy guaranteed by Lemma 9.5. In particular, the supernode selected by the duplicator has the same supercolor as the supernode selected by the spoiler. As for his selection of a point "within" the supernode, there are two cases. In the first case, the graphs corresponding to both supernodes are the same (either both G_0 or both G_1), and the duplicator makes the identity move: he selects the same point in the graph G_0 or G_1 as the spoiler did. In the other case (one supernode corresponds to G_0 and the other to G_1), the duplicator's choice is based on his winning strategy in the symmetric (c, r_2)-game over G_0, G_1.

We leave to the reader the straightforward verification that we have described an overall winning strategy for the duplicator. ∎

We now show that property $\mathscr{P}_2$ belongs to closed monadic NP. We first need some preliminary results showing that two other undirected graph properties belong to closed monadic NP.

The first preliminary step is to show that the property "Each inner factor is connected" belongs to the positive first-order closure of monadic NP, and therefore to closed monadic NP. This result is interesting in its own right, since on the face of it, it seems that this property should require at least one universal set (unary second-order) quantifier. Nonetheless, we show that existential set quantifiers suffice, provided that we can place some first-order quantifiers in front. (It should perhaps be noted that this property does not belong to monadic NP, without first-order quantifiers in front. This follows from Theorems 5.2 and 7.1, since (1) the only inner factors of a single cycle of length $n \geqslant 5$ are the cycle itself and single points, and (2) any nonconnected graph, such as two disjoint cycles, contains a nonconnected inner factor, namely, the graph itself.)

PROPOSITION 9.7. *The property "Each inner factor is connected" belongs to the positive first-order closure of monadic NP.*

Proof. It is more natural to argue that the complementary property, "There exists a nonconnected inner factor", belongs to the positive first-order closure of monadic co-NP, that is, can be expressed by a sentence of the form $\mathbf{PR}\varphi$, where $\mathbf{P}$ consists of first-order quantifiers, $\mathbf{R}$ consists of universal unary second-order quantifiers, and φ is first-order.

If C is a set of points of a graph, a *C-free path* is a path in the graph that contains no point of C. If x and y are points of a graph and C is a set of points, we say that x and y are *C-free path-connected* if there is a C-free path between x and y; in particular, neither x nor y belongs to C. Let

$$Common(x, y) = \{w \mid w \neq x, w \neq y, \text{ and } (w, x) \text{ and } (w, y) \text{ are edges}\}.$$

Thus, $Common(x, y)$ contains all points, other than x and y, that have an edge with both x and y.

We argue that the property "There exists a nonconnected inner factor" is expressed by the following sentence, where x, y, z denote points and C denotes a set of points:

$\exists x\, \exists y\, \forall z\, \forall C$

if $C = Common(x, y)$ then:

(1) if z is C-free path-connected to either x or y, then there are edges between z and every point of C, and

(2) x and y are not C-free path-connected.

(C can be eliminated from the sentence since $Common(x, y)$ has a first-order definition. But it is more convenient to argue about the correctness of the sentence with C present.) It is easy to see that this sentence can be transformed, by standard quantifier manipulation, to one of the form required for the positive first-order closure of monadic co-NP. It is enough to note that "s and t are C-free path-connected" can be expressed by a monadic NP formula. Such a formula is $\exists A((A \cap C = \emptyset) \wedge path(A, s, t))$, where $path(A, s, t)$ is the first-order formula described in Section 4 that says that A is a strict path between s and t. So "s and t are not C-free path-connected" can be expressed by a monadic co-NP formula. In transforming the sentence, we use that, if ψ_1 is a monadic NP formula and ψ_2 is a monadic co-NP formula, then $(\psi_1 \Rightarrow \psi_2)$ is equivalent to a monadic co-NP formula.

It remains to show that this sentence expresses the desired property. Say first that the graph G has a nonconnected inner factor H. Let M be the set of points of H (so M is a module), and let A and B be a partition of M such that A and B are nonempty and there is no edge between a point of A and a point of B. To show that G satisfies the sentence, choose $x \in A$ and $y \in B$ (arbitrarily). Let $C = Common(x, y)$ since this is the only instantiation of C that is relevant to the truth of the sentence. The following claim will be useful.

Claim. If $w \in M$ and $u \notin M$, then w and u are not C-free path-connected.

Proof of Claim. Assume for contradiction that there is a path $v_1, v_2, ..., v_k$ of points such that $v_1 = w$, $v_k = u$, there is an edge between v_i and v_{i+1} for $1 \leqslant i < k$, and $v_i \notin C$ for all i. Since $w \in M$ and $u \notin M$, there must be some i with $1 \leqslant i < k$ such that $v_i \in M$ and $v_{i+1} \notin M$. Since v_{i+1} is not in M and has an edge with $v_i \in M$, the definition of module implies that v_{i+1} has an edge with every point of M, including x and y. Also, $v_{i+1} \neq x$ and $v_{i+1} \neq y$ because $v_{i+1} \notin M$. So $v_{i+1} \in Common(x, y) = C$, contradicting the fact that the path is C-free. This completes the proof of the claim. ∎

Using the claim, we can now show that (1) and (2) hold. We first note that $C \cap M = \emptyset$. For if there were a point $w \in C \cap M$, then the edges (x, w) and (w, y) would form a path lying within M, contradicting that x and y belong to different parts of the partition of M into A and B. To see that (1) holds for all z, let z be C-free path-connected to either x or y. Since $x, y \in M$, the claim implies that $z \in M$. By the definition of module, z has an edge with every point of C, since every point of C does not belong to M and has an edge with a member x of M. To see that (2) holds, suppose that x and y are C-free path-connected. Since x and y belong to different parts of the partition of M into A and B, each path connecting x and y must visit a point $u \notin M$. But by the claim, this path cannot be C-free.

Now say that the sentence is satisfied by a graph G. Let x and y be such that the part of the sentence following $\exists x \, \exists y$ is satisfied. Let $C = Common(x, y)$. Let M be the set of points of G that are C-free path-connected to either x or y. In particular,

$x, y \in M$ and $M \cap C = \varnothing$. We argue that M is a module, and *ind*(M) (the subgraph of G induced by M) is nonconnected.

To show that M is a module, let $v_1, v_2 \in M$ and $u \notin M$ be arbitrary. Assume that there is an edge (u, v_1); we must show that there is an edge (u, v_2). We first argue that $u \in C$. Suppose not, i.e., that $u \notin C$. Now there is an edge between u and v_1, and by definition of M there is a C-free path between v_1 and either x or y. So there is a C-free path between u and either x or y, contradicting the choice of u that $u \notin M$. So we have $u \in C$. Since $v_2 \in M$, the definition of M and part (1) of the formula imply that v_2 has an edge with every point of C. In particular, there is an edge (u, v_2), which was to be shown.

That *ind*(M) is nonconnected follows from part (2) of the formula (x and y are not C-free path-connected) and the fact noted above that $M \cap C = \varnothing$. For if x were connected to y by a path that stays within M, this would be a C-free path. ▌

The next preliminary step is to show that the complement of property $\mathscr{P}_1$ belongs to closed monadic NP. This follows easily from Proposition 9.7, as we now show.

PROPOSITION 9.8. *The property "At least two connected components have each of their inner factors connected" belongs to closed monadic NP.*

Proof. Clearly, "At least two connected components have each of their inner factors connected" is expressed by the following:

There exist sets A and B of points such that:

(1) A and B are nonempty and disjoint;

(2) both *ind*(A) and *ind*(B) are connected components; and

(3) both *ind*(A) and *ind*(B) have each of their inner factors connected.

To finish the proof, we show that (2) and (3) can be expressed by closed monadic NP sentences. (Clearly, (1) is first-order expressible.) That *ind*(X) is a connected component can be expressed by saying: (i) x and y are path-connected for every x and y in X; and (ii) there is no edge between a point in X and point not in X. That *ind*(X) has each of its inner factors connected can be expressed by a sentence as in the proof of Proposition 9.7 by restricting everything to points in X. ▌

We can now give the following upper bound.

THEOREM 9.9. *Property $\mathscr{P}_2$ belongs to closed monadic NP.*

Proof. As we now show, a sentence expressing that G has property $\mathscr{P}_2$ has the form

$$\exists A\, \exists W\, \exists P(\psi_1 \wedge \psi_2 \wedge \psi_3 \wedge \psi_4)$$

where A, W, P are unary relation symbols and each ψ_i is a closed monadic NP sentence. We now explain what these relations mean in the construction, and the constraints that are placed on them by the ψ_i's.

The sentence ψ_1 says that A represents a set of points such that *ind*(A) is a cycle of length at least five. In more detail, ψ_1 says that A contains at least five points, every point of A has an edge with exactly two other points of A, and $ind(A)$ is connected, i.e., every two points of A are path-connected by a path that stays in A. We have noted above that path-connectivity can be expressed by a monadic NP sentence, so ψ_1 can be a monadic NP sentence. For $x \in A$, the *neighbors* of x are the two points of A that have an edge with x.

Before describing ψ_2, we must define some sets that depend on the choice of A. For $x \in A$, let M_x be the set of points that have an edge with both of the neighbors of x; in particular, $x \in M_x$. The first-order sentence ψ_2 places the following constraints on the sets M_x:

1. each point of G is in M_x for exactly one x, and

2. for all $x, y \in A$ and all u, w, if $x \neq y$ and $u \in M_x$ and $w \in M_y$, then there is an edge between u and w iff there is an edge between x and y.

Assume that ψ_1 and ψ_2 hold. Let $x_1, ..., x_n$ be the points of A. For $1 \leqslant i \leqslant n$, let H_i be the subgraph induced by M_{x_i}. Let C be the cycle induced by A. It should be evident that $G = C[H_1, ..., H_n]$. Conversely, if $G = C[H_1, ..., H_n]$ for some cycle C and graphs $H_1, ..., H_n$, then there is a choice of A that satisfies ψ_1 and ψ_2 and, for some naming $x_1, ..., x_n$ of the points in A, the set M_{x_i} is the set of points of H_i, for all i. Viewing G as C with the graphs H_i substituted for the points of C, let A contain one (arbitrary) point from each H_i.

The relation W gives a color to each point x, either white if Wx, or black if $\neg Wx$. The sentence ψ_3 says that, for all x, the point x is colored white iff $x \in A$ and $ind(M_x)$ has property $\mathscr{P}_1$. Recall that we have shown in Theorem 8.4 and Proposition 9.8 that both $\mathscr{P}_1$ and its complement belong to closed monadic NP. Also, there is a first-order formula that says "$u \in M_x$". Therefore, it is easy to see that ψ_3 can be a closed monadic NP sentence.

Finally, the relation P and the formula ψ_4 are used to check that A contains an even number of white points. (All points outside of A are black.) Think of P as giving an additional two-coloring to the white points. Say that distinct white points x and y of A are *adjacent* if there is a path in A between x and y that visits no white point other than x and y. The formula ψ_4 is designed to say that every two adjacent white points are colored differently by P. Clearly, such a coloring by P exists iff the number of white points is even. Let N_W denote the number of white points. The formula ψ_4 says that either $N_W = 0$, or

$$N_W \geqslant 2 \wedge \forall x \, \forall y (\text{“}x \text{ and } y \text{ are not adjacent white points”} \vee (\mathrm{P}x \Leftrightarrow \neg \mathrm{P}y)).$$

Assuming that A contains at least 2 white points, the property "x and y are not adjacent white points" can be expressed by a monadic NP formula. This formula says that either x is not white, or y is not white, or $x = y$, or there exist sets A_1 and A_2 such that $A_1 \cup A_2 = A$, $A_1 \cap A_2 = \{x, y\}$, both A_1 and A_2 contain a white point other than x and y, and $path(A_1, x, y)$ and $path(A_2, x, y)$. So ψ_4 can be a closed monadic NP formula. ∎

The following theorem is an immediate consequence of Theorems 9.6 and 9.9.

THEOREM 9.10. *Property $\mathcal{P}_2$ is in closed monadic NP but not in the first-order closure of monadic NP.*

This theorem immediately implies Theorem 1.2, which is one of our main separation results.

10. CLOSED MONADIC NP VERSUS THE MONADIC HIERARCHY

In this section we give another indication of the expressive power of closed monadic NP, by showing that closed monadic NP contains properties at arbitrarily high levels of the monadic hierarchy. Just as monadic NP is a restricted version of NP, the monadic polynomial-time hierarchy (called the "monadic hierarchy" for short) is a restricted version of the polynomial-time hierarchy [Sto77]. A Σ^1_k formula is a formula of the form $Q_1 Q_2 \cdots Q_k \varphi$, where Q_i consists of existential (resp., universal) second-order quantifiers if i is odd (resp., even), and φ is first-order. Following Fagin's [Fag74] proof that the properties in NP are precisely the properties expressible by Σ^1_1 sentences, Stockmeyer [Sto77] observed that the properties in the class Σ^{P}_k of the polynomial-time hierarchy are precisely the properties expressible by Σ^1_k sentences. A *monadic Σ^{P}_k formula* is a monadic Σ^1_k formula, that is, a Σ^1_k formula where all second-order quantifiers quantify only over unary relations (i.e., sets). Let *monadic Σ^{P}_k* be the class of properties expressible by a monadic Σ^{P}_k sentence. Similarly, the class *monadic Π^{P}_k* is defined in terms of sentences $Q_1 Q_2 \cdots Q_k \varphi$, where Q_i consists of universal (resp., existential) unary second-order quantifiers if i is odd (resp., even), and φ is first-order. The *monadic (polynomial-time) hierarchy* is the union of the monadic Σ^{P}_k classes over all k. It is clear that closed monadic NP is contained in the monadic hierarchy, because every closed monadic NP sentence $\mathbf{Q}\varphi$, where φ is first-order, can be transformed to an equivalent monadic Σ^{P}_k sentence $\mathbf{Q}'\varphi'$, for some k (depending on $\mathbf{Q}$): Each first-order quantifier in $\mathbf{Q}$ is replaced by a unary second-order quantifier in $\mathbf{Q}'$, and φ is modified to restrict all newly introduced second-order quantifiers to quantify only over singleton sets. Since the k in this transformation depends on $\mathbf{Q}$, this raises the question of whether there is a fixed k such that closed monadic NP is contained in monadic Σ^{P}_k. In this section, we show that the answer is "no": For each k, there is a property of undirected graphs that belongs to closed monadic NP, but does not belong to monadic Σ^{P}_k.

Fagin [Fag94] asked whether the monadic hierarchy is strict, i.e., whether monadic Σ^{P}_k does not equal monadic $\Sigma^{\mathrm{P}}_{k+1}$ for all $k \geqslant 1$. Matz and Thomas [MT97b] have shown that the monadic hierarchy is strict over graph properties. (Interestingly, Matz and Thomas's proof does not use Ehrenfeucht–Fraïssé games, but is automata-theoretic and arithmetical in nature. Their proof method is described in Section 12.) Building on the work of Matz and Thomas, Schweikardt [Sch97] shows that, for each $k \geqslant 1$, there is a property in both monadic $\Sigma^{\mathrm{P}}_{k+1}$ and monadic Π^{P}_{k+1} but not in the Boolean closure of monadic Σ^{P}_k. See also [MST]. As a step toward their result, Matz and Thomas prove that certain types of sentences involving a transitive closure operator can express properties arbitrarily high in the monadic hierarchy; this is useful to us in proving the same thing about the expressive power of closed monadic NP sentences.

Following [MT97b], we define the class of $\mathrm{FO}^{TC(1)}$ formulas (first-order formulas with transitive closure). Every first-order formula is a $\mathrm{FO}^{TC(1)}$ formula, and the $\mathrm{FO}^{TC(1)}$ formulas are closed under Boolean operations and first-order quantification. In addition, if $\varphi(u, u', \vec{z})$ is a $\mathrm{FO}^{TC(1)}$ formula in which the first-order variables $u, u', \vec{z}$ occur freely, where $\vec{z}$ abbreviates a sequence $z_1, ..., z_k$ of first-order variables for some $k \geqslant 0$, then $TC(u, u', \varphi)(x, y, \vec{z})$ is a $\mathrm{FO}^{TC(1)}$ formula in which $x, y, \vec{z}$ are free and u, u' are bound. For S a suitable structure, $S \models TC(u, u', \varphi)$ $(x, y, \vec{z})$ iff there are points $u_0, u_1, ..., u_n$ (for some $n \geqslant 0$) such that $x = u_0$, $y = u_n$, and $S \models \varphi(u_i, u_{i+1}, \vec{z})$ for all i with $0 \leqslant i < n$. A *monadic* $\Sigma_1^{TC(1)}$ *formula* is a formula of the form $\mathbf{R}\varphi$, where $\mathbf{R}$ consists of existential unary second-order quantifiers, and φ is a $\mathrm{FO}^{TC(1)}$ formula. (Matz and Thomas call a formula of this form simply a $\Sigma_1^{TC(1)}$ formula. We attach the adjective "monadic" to be consistent with our other terminology. The "1" in $TC(1)$ in these notations means that formulas contain a "1-dimensional" transitive closure operator.)

The result of Matz and Thomas that we use is the following. (See [MST] for the proof of a stronger result,)

THEOREM 10.1. *For each* $k \geqslant 1$, *there is a property* $\mathcal{S}$ *of directed graphs such that* $\mathcal{S}$ *is expressible by some monadic* $\Sigma_1^{TC(1)}$ *sentence, but* $\mathcal{S}$ *does not belong to monadic* Σ_k^{P}.

To translate this to a result about closed monadic NP, we show the following.

PROPOSITION 10.2. *If* σ *is a sentence in the closure of the* $\mathrm{FO}^{TC(1)}$ *formulas under first-order quantification and existential unary second-order quantification, then* σ *is equivalent to a closed monadic NP sentence.*

Proof. Clearly, it is enough to show that every $\mathrm{FO}^{TC(1)}$ formula is equivalent to a closed monadic NP formula, because the class of closed monadic NP formulas is closed under first-order quantification and existential unary second-order quantification.

The key observation is that the transitive closure operator is closely related to the directed reachability property considered in Section 4. Specifically, $TC(u, u', \varphi)$ $(x, y, \vec{z})$ is satisfied by a structure S iff there is a directed path from x to y in the directed graph whose set of points is the universe of S and whose edge relation E is defined by $Euu' = \varphi(u, u', \vec{z})$. We have shown in Theorem 4.1 that there is a closed monadic NP formula ρ that says that "There is a directed path from s to t in the graph with edge relation E". Write ρ in the form $\mathbf{Q}(\bigvee_{i=1}^{m_1} \bigwedge_{j=1}^{m_2} \rho_{i,j})$, where $\mathbf{Q}$ contains first-order quantifiers and existential unary second-order quantifiers, and each $\rho_{i,j}$ is an atomic formula or its negation. An atomic formula is a formula containing no logical connectives or quantifiers; thus, if the language contains only the binary relation symbol E, then the atomic formulas are of the form Evv', Pv, and $v = v'$, where v, v' are variables, and P is a unary relation symbol contained in $\mathbf{Q}$. From this it is clear that, if both φ and $\neg\varphi$ are equivalent to closed monadic NP formulas, then $TC(u, u', \varphi)(x, y, \vec{z})$ is equivalent to a closed monadic NP formula. In addition, we shall need the same result for $\neg TC(u, u', \varphi)(x, y, \vec{z})$. It is enough to note that there is a monadic NP formula $\bar{\rho}$ that says that "There is no directed path from s to t in the graph with edge relation E": The formula $\bar{\rho}$ says that there is a set A of points such that $s \in A$, $t \in \bar{A}$, and there is no edge directed from a point of A to a point of $\bar{A}$. Thus, if both φ and $\neg\varphi$ are equivalent to closed monadic NP formulas, then both $TC(u, u', \varphi)(x, y, \vec{z})$ and $\neg TC(u, u', \varphi)(x, y, \vec{z})$ are equivalent to closed monadic NP formulas.

It is now straightforward to prove that every $\mathrm{FO}^{TC(1)}$ formula is equivalent to some closed monadic NP formula, by induction on the depth of nesting of the TC operators in ψ. The *transitive closure depth* $\mathrm{TCD}(\psi)$ of a $\mathrm{FO}^{TC(1)}$ formula ψ is defined recursively as follows: $\mathrm{TCD}(\psi) = 0$ if ψ is first-order; $\mathrm{TCD}(\neg\psi) = \mathrm{TCD}(\psi)$; $\mathrm{TCD}(\psi_1 \wedge \psi_2) = \max\{\mathrm{TCD}(\psi_1), \mathrm{TCD}(\psi_2)\}$; $\mathrm{TCD}(\exists\psi) = \mathrm{TCD}(\psi)$; $\mathrm{TCD}(TC(u, u', \psi)(x, y, \vec{z})) = 1 + \mathrm{TCD}(\psi)$.

Clearly, a $\mathrm{FO}^{TC(1)}$ formula ψ is equivalent to a closed monadic NP formula if $\mathrm{TCD}(\psi) = 0$, so let ψ be a $\mathrm{FO}^{TC(1)}$ formula with $\mathrm{TCD}(\psi) = d \geqslant 1$. We can assume that ψ is of the form $\mathbf{P}(\bigvee_{i=1}^{m_1} \bigwedge_{j=1}^{m_2} \psi_{i,j})$, where $\mathbf{P}$ contains first-order quantifiers, and for each i, j, either $\psi_{i,j}$ is first-order or $\psi_{i,j}$ has the form $TC(u, u', \varphi_{i,j})(x, y, \vec{z})$ or the form $\neg TC(u, u', \varphi_{i,j})(x, y, \vec{z})$ for some $\mathrm{FO}^{TC(1)}$ formula $\varphi_{i,j}$ with $\mathrm{TCD}(\varphi_{i,j}) < d$; therefore, also $\mathrm{TCD}(\neg\varphi_{i,j}) < d$. By the induction hypothesis, both $\varphi_{i,j}$ and $\neg\varphi_{i,j}$ are equivalent to closed monadic NP formulas. Therefore, by the argument above, both $TC(u, u', \varphi_{i,j})(x, y, \vec{z})$ and $\neg TC(u, u', \varphi_{i,j})(x, y, \vec{z})$ are equivalent to closed monadic NP formulas. Thus, in ψ we can replace each $\psi_{i,j}$ by a closed monadic NP formula, so ψ itself is equivalent to a closed monadic NP formula. ∎

It follows immediately from Theorem 10.1 and Proposition 10.2 that for each $k \geqslant 1$ there is a directed graph property that is in closed monadic NP but not in monadic Σ_k^{P}. By encoding directed graphs as undirected graphs (as in [MST]), it follows further that there is an undirected graph property that is in closed monadic NP but not in monadic Σ_k^{P}. We thus obtain the following result, which is a slight restatement of Theorem 1.3 in the introduction.

THEOREM 10.3. *For each $k \geqslant 1$, there is an undirected graph property that is in closed monadic NP but not in monadic Σ_k^{P}.*

Theorem 10.3 was shown independently by Matz and Thomas [MT97a] (see [MST] for this and more recent improvements). Their argument does not require Proposition 10.2, since it is based on the fact that transitive closure is used in a restricted way in the monadic $\Sigma_1^{TC(1)}$ sentences constructed to prove Theorem 10.1. Even though Proposition 10.2 is not needed to obtain Theorem 10.3 if one knows the details of the proof of Theorem 10.1, we feel that Proposition 10.2 is interesting in its own right as an indication of the power of closed monadic NP.

On the other hand, assuming that NP $\neq$ co-NP, there is a property in monadic co-NP (and, therefore, in monadic Σ_k^{P} for all $k \geqslant 2$), but not in closed monadic NP. An example of such a property is non-3-colorability. We noted in Section 2 that this property belongs to monadic co-NP. Since non-3-colorability is co-NP-complete [GJS76] and since closed monadic NP is contained in NP, if non-3-colorability belonged to closed monadic NP then there would be a co-NP-complete property in NP, contrary to the assumption that NP ≠ co-NP. To summarize, it follows from the assumption that NP ≠ co-NP that there is a property in the monadic hierarchy (and in particular, in monadic co-NP) but not in closed monadic NP. It is presently open, however, whether there is a property in the monadic hierarchy but not in closed monadic NP, without making an unproved complexity-theoretic assumption.

11. THE CLOSED MONADIC HIERARCHY

In the definition of closed monadic NP, we focused on the configuration of the second-order quantifiers. In the same spirit, we can define a monadic version of Σ_k^{P} where we focus on the pattern of the second-order quantifiers, and allow arbitrary interleavings of first-order quantifiers among the unary second-order quantifiers. Thus, let us define closed monadic Σ_k^{P}, which bears the same relationship to monadic Σ_k^{P} as closed monadic NP bears to monadic NP. For integer $k \geqslant 1$, define a *closed monadic* Σ_k^{P} *formula* to be a formula of the form $\mathbf{Q}\varphi$, where (a) the prefix $\mathbf{Q}$ can contain an interleaving of unary second-order quantifiers and first-order quantifiers, (b) the leading second-order quantifier is existential, (c) when we ignore the first-order quantifiers, there are at most $k-1$ alternations between existential and universal second-order quantifiers, and (d) φ is first-order. Let *closed monadic* Σ_k^{P} be the class of properties expressible by a closed monadic Σ_k^{P} sentence. In particular, closed monadic Σ_1^{P} is precisely closed monadic NP. The *closed monadic hierarchy* is the union of closed monadic Σ_k^{P} classes over all k. Clearly, for all k, monadic Σ_k^{P} is contained in closed monadic Σ_k^{P}; and every property in the closed monadic hierarchy is also in the monadic hierarchy (possibly at a higher level). However, as noted in Section 10, it is an open question whether the entire monadic hierarchy is contained in closed monadic Σ_1^{P}. Thus, it is an open question whether the closed monadic hierarchy collapses to the first level. Generally, it is an open question whether the closed monadic hierarchy is strict, that is, whether closed monadic $\Sigma_{k+1}^{\mathrm{P}}$ does not equal closed monadic Σ_k^{P} for all $k \geqslant 1$. (We note that there are obstacles to settling these questions using the same method that Matz and Thomas [MT97b] used to prove strictness of the monadic hierarchy. We discuss this further in Section 12.) In this section we show that if the polynomial-time hierarchy is strict then the closed monadic hierarchy is strict.

We first note an upper bound on the complexity of properties in closed monadic Σ_k^{P}.

PROPOSITION 11.1. *For all* $k \geqslant 1$, *closed monadic* Σ_k^{P} *is contained in* Σ_k^{P}.

Proof. Let $\mathcal{S}$ be a property in closed monadic Σ_k^{P}. Therefore, $\mathcal{S}$ is expressible by a closed monadic Σ_k^{P} sentence. We show that each closed monadic Σ_k^{P} formula is equivalent to a Σ_k^1 formula; that is, we can move all first-order quantifiers to the right of all second-order quantifiers, while possibly increasing the arity of the second-order quantifiers, but without changing the number of alternations of second-order quantifiers. All first-order quantifiers are moved to the right by repeatedly applying the following replacements. Let x denote a first-order variable and R denote a relation symbol. Clearly $\exists x \exists R$ can be replaced by $\exists R \exists x$, and $\forall x \forall R$ can be replaced by $\forall R \forall x$. The other two cases require an increase in the arity of the relation. If R represents an r-ary relation, then $\forall x \exists R$ is replaced by $\exists S \forall x$, where S represents an $(r+1)$-ary relation. The r-ary relation R that corresponds to x is the set of all $\langle y_1, ..., y_r \rangle$ such that $\langle x, y_1, ..., y_r \rangle$ is in S. In the first-order part of the sentence, we simply replace R by S. The last case is the logical dual of the preceding one: $\exists x \forall R$, where R is r-ary, is replaced by $\forall S \exists x$, where S is $(r+1)$-ary. Thus, $\mathscr{S}$ is expressible by a Σ_k^1 sentence. But it is known [Sto77] that the properties expressible by Σ_k^1 sentences are precisely the properties in Σ_k^{P}. ∎

We noted in Section 2 that monadic NP contains an undirected graph property that is NP-complete. The next result generalizes this to the monadic hierarchy.

THEOREM 11.2. *For all $k \geqslant 1$, there is an undirected graph property that belongs to monadic Σ_k^{P} and is Σ_k^{P}-complete.*

Before proving this result, we show how it is used to obtain the main result of this section, concerning strictness of the closed monadic hierarchy.

THEOREM 11.3. *Assume $k \geqslant 1$. If $\Sigma_{k+1}^{\mathrm{P}} \neq \Sigma_k^{\mathrm{P}}$ then there is an undirected graph property that belongs to monadic $\Sigma_{k+1}^{\mathrm{P}}$ but does not belong to closed monadic Σ_k^{P}. In particular, if the polynomial-time hierarchy is strict, then so is the closed monadic hierarchy.*

Proof. Let $\mathscr{S}$ be the undirected graph property, guaranteed by Theorem 11.2, that belongs to monadic $\Sigma_{k+1}^{\mathrm{P}}$ and is $\Sigma_{k+1}^{\mathrm{P}}$-complete. Suppose for contradiction that $\mathscr{S}$ belongs to closed monadic Σ_k^{P}. By Proposition 11.1, we would then have a $\Sigma_{k+1}^{\mathrm{P}}$-complete property in Σ_k^{P}, contradicting the assumption that $\Sigma_{k+1}^{\mathrm{P}} \neq \Sigma_k^{\mathrm{P}}$. ∎

In the remainder of this section we prove Theorem 11.2. One way to see that monadic Σ_k^{P} contains a Σ_k^{P}-complete property is to take a known Σ_k^{P}-complete problem, such as the quantified Boolean formula problem restricted to k blocks of alternating quantifiers, proved Σ_k^{P}-complete in [Sto77, Wra77], and encode instances of the problem as structures in such a way that the encoded problem belongs to monadic Σ_k^{P}. An advantage of this approach is that one does not have to show that the problem is Σ_k^{P}-complete. This is the approach of Makowsky and Pnueli [MP96] (done independently of us) who encode quantified Boolean formulas as structures over ordered universes. However, encoding the quantified Boolean formula problem as a graph property over unordered universes leads to a graph property with a somewhat unnatural definition. Our approach is to obtain the result for a more natural undirected graph property, a generalization of 3-colorability. Although we must then show that the new property is Σ_k^{P}-complete, a benefit is that we get a new Σ_k^{P}-complete property, which may be of independent interest.

The undirected graph property that we consider is a generalization of 3-colorability to a k-round game between two players, player 1 and player 2. We let

1 and 2 denote the integers mod 2, so $i \bmod 2$ is 1 (resp., 2) if i is odd (resp., even). A graph G has the property *k-round* 3-*colorability* if player 1 has a winning strategy in the following k-round game:

- At round i for $1 \leqslant i < k$, player $i \bmod 2$ assigns a color from the set $\{1, 2, 3\}$ to every vertex of the graph having degree i.

- At round k, player $k \bmod 2$ assigns a color from the set $\{1, 2, 3\}$ to every vertex that has not been colored in a previous round, i.e., every vertex whose degree is at least k.

- The player who moved last (player $k \bmod 2$) wins the game if the entire coloring is legal, i.e., if every pair of distinct vertices that are connected by an edge are colored differently.

In particular, a graph is 1-round 3-colorable iff it is 3-colorable in the usual sense.

In the definition of this property, the degrees of vertices are used to partition the vertices into k sets, where each set contains vertices that are colored during the same round. Another possibility would be to define a *modified k-round* 3-*coloring game* where an instance of the game consists of a graph together with a partition of the vertices into k sets $V_1, ..., V_k$; the vertices in the set V_i are colored by player $i \bmod 2$ during round i, and the winner is determined as in the original k-round 3-coloring game defined above. A drawback of the modified game is that an instance of the game must somehow specify the vertex partition. For example, the vertex partition could be specified by k additional unary relations, but then the property would be a property of a structure over one binary relation and several unary relations. Alternatively, the partition could be specified by attaching different types of graph "gadgets" to the vertices, but this would complicate the definition of the property. By using degrees to specify the sets, the signature of the property is kept simple (a single binary relation) and the definition of the property is kept simple as well. We show that k-round 3-colorability is Σ_k^{P}-complete even when the degree of the graph is restricted to $\max\{k, 4\}$. We later note that this degree restriction cannot be lowered, except possibly for $k = 2$, 3, assuming that the polynomial-time hierarchy does not collapse. The restriction on maximum degree is not needed to obtain Theorems 11.2 and 11.3, but we include the further information about degree restriction since it may be of independent interest, as a result about a new Σ_k^{P}-complete problem.

Theorem 11.2 is an immediate consequence of the following.

THEOREM 11.4. *For each $k \geqslant 1$, k-round 3-colorability belongs to monadic Σ_k^{P} and is Σ_k^{P}-complete. Moreover, k-round 3-colorability is Σ_k^{P}-complete even when restricted to graphs of maximum degree at most* $\max\{k, 4\}$.

Proof. We show here that k-round 3-colorability belongs to monadic Σ_k^{P}. Since the details of the polynomial-time reduction showing Σ_k^{P}-completeness lie somewhat outside the main theme of this paper, description of this reduction has been moved to the appendix.

We describe a monadic Σ_k^{P} sentence that expresses k-round 3-colorability. The sentence contains unary relation symbols A_{ij} for $1 \leqslant i \leqslant k$ and $1 \leqslant j \leqslant 3$. For each i, certain interpretations of the three relation symbols A_{i1}, A_{i2}, A_{i3}, which we call *proper interpretations*, encode the colors of vertices that are supposed to be colored during round i. An interpretation of A_{i1}, A_{i2}, A_{i3} is *proper* if (i) for each vertex u that is supposed to be colored during round i, exactly one of A_{i1}, A_{i2}, A_{i3} is true (where A_{ij} true means that u is colored j), and (ii) for each vertex u that is not supposed to be colored during round i, the relations A_{i1}, A_{i2}, A_{i3} are all false. Since the round at which a vertex is supposed to be colored is determined by its degree (either $1, 2, 3, ..., k-1$, or at least k), and since for each d there are first-order formulas that express "the degree of vertex u is d" and express "the degree of vertex u is at least d", it is clear that, for each i, there is a first-order formula $proper_i$ that expresses "the interpretation of A_{i1}, A_{i2}, A_{i3} is proper".

Let E be the edge relation. Let *legal* be a first-order formula expressing that the coloring encoded by the relations A_{ij} is legal, assuming that the interpretation of these relations is proper; that is, *legal* is

$$\forall u \, \forall v \left(\bigwedge_{1 \leqslant j \leqslant 3} \left(\bigvee_{1 \leqslant i \leqslant k} A_{ij} u \wedge \bigvee_{1 \leqslant i \leqslant k} A_{ij} v \right) \Rightarrow \neg Euv \right).$$

The sentence expressing k-round 3-colorability has the form

$$\exists A_{11} \, \exists A_{12} \, \exists A_{13} \, \forall A_{21} \, \forall A_{22} \, \forall A_{23} \, \exists A_{31} \, \exists A_{32} \, \exists A_{33} \cdots Q_k A_{k1} Q_k A_{k2} Q_k A_{k3} \psi$$

where Q_k is $\exists$ ($\forall$) if k is odd (even), and ψ is first-order. Recall that in order for player 1 to win the game, player 1 wants the final coloring to be legal (resp., not legal) if k is odd (resp., even). First-order formulas ψ_i are defined by reverse induction on i for $1 \leqslant i \leqslant k+1$. First, ψ_{k+1} is *legal* if k is odd, or $\neg$*legal* if k is even. The definition of the other ψ_i's are intended to restrict all second-order quantification to proper interpretations. Recalling that the relation symbols A_{i1}, A_{i2}, A_{i3} are existentially (resp., universally) quantified if i is odd (resp., even), we see this restriction is accomplished by defining

$$\psi_i \equiv \begin{cases} (proper_i \wedge \psi_{i+1}) & \text{if } i \text{ is odd} \\ (proper_i \Rightarrow \psi_{i+1}) & \text{if } i \text{ is even.} \end{cases}$$

Then, ψ is ψ_1. This completes the demonstration that k-round 3-colorability belongs to monadic Σ_k^{P}. ∎

Remark. We now discuss the optimality of the degree restriction $\max\{k, 4\}$, under the assumption that the polynomial-time hierarchy does not collapse. Let (k, d)-*colorability* abbreviate k-round 3-colorability restricted to graphs of maximum degree at most d. It is obvious that (k, d)-colorability belongs to Σ_d^{P}, because there are at most d rounds during which some vertex is actually colored. Therefore, assuming that the polynomial-time hierarchy does not collapse, (k, d)-colorability is Σ_k^{P}-complete only if $k \leqslant d$. In other words, the first term in $\max\{k, 4\}$ cannot be reduced below k (if $k \geqslant 5$). Two observations are relevant toward reducing the second term, 4, when $k \leqslant 3$. First, if $k=1$, it is known that $(1, 3)$-colorability (ordinary 3-colorability restricted to graphs of maximum degree at most 3) belongs to P [GJS76]. Second, if $k=2$, it is easy to see that $(2, 2)$-colorability belongs to P, since player 1 wins iff the graph contains two vertices of degree 1 that are connected by an edge. This leaves only two cases where it is not clear whether (k, d)-colorability is Σ_k^{P}-complete, namely $(k, d) = (2, 3), (3, 3)$. We have not seriously tried to settle these cases.

12. GROWTH ARGUMENTS

The method by which Matz and Thomas [MT97b] prove strictness of the monadic hierarchy is based on the growth rates of functions definable by sentences (for a certain notion of function definability, described below). A "growth argument" was used earlier by Otto [Ott95] to show that increasing the number of second-order quantifiers in a monadic NP sentence allows more properties to be expressed. In this section we note that a version of the main result of Section 9, that there is a property in closed monadic NP but not in the first-order closure of monadic NP, can be proved by a growth argument. Growth arguments are done in the context of *grid structures*; a grid structure consists of a rectangular array of points and two successor relations, a vertical successor and a horizontal successor. As shown in [MT97b, MST], results obtained for grid properties can often be translated to similar results for graph properties, by encoding grids as graphs.

An (m, n)-grid is a $\{S_1, S_2\}$-structure, where S_1, S_2 are binary relations, where the set of points is an $m \times n$ array (the points are $\langle i, j \rangle$ for $1 \leqslant i \leqslant m$ and $1 \leqslant j \leqslant n$), and S_1 is the vertical successor relation ($\langle \langle i, j \rangle, \langle i+1, j \rangle \rangle \in S_1$), and S_2 is the horizontal successor relation ($\langle \langle i, j \rangle, \langle i, j+1 \rangle \rangle \in S_2$). If φ is a $\{S_1, S_2\}$-sentence and f is a function from positive integers to positive integers, we say that φ *defines* f if for each $m \geqslant 1$ there is exactly one n, namely $n = f(m)$, such that the (m, n)-grid satisfies φ.

Define the iterated exponential functions s_k by $s_0(x) = x$, and $s_{k+1}(x) = 2^{s_k(x)}$. Say that a function f is *at most k-fold exponential* if $f(m) \leqslant s_k(cm)$ for some constant $c > 0$ and for every m. Say that f is *at least k-fold exponential* if $f(m) \geqslant s_k(cm)$ for some constant $c > 0$ and for every m.

Before giving the main result of this section, we outline how growth arguments were used to prove strictness of the monadic hierarchy.

THEOREM 12.1 [MT97b]. *For all* $k \geqslant 1$:

(1) *There is a function that is at least k-fold exponential and that can be defined by a monadic* $\Sigma_{2k+3}^{\mathrm{P}}$ *sentence.*

(2) *If the function f can be defined by a Boolean combination of monadic* Σ_k^{P} *sentences, then f is at most k-fold exponential.*

It follows from Theorem 12.1 that the monadic hierarchy (over grid properties) does not collapse to any fixed level. A refinement of the argument, given in [MT97b], shows that the monadic hierarchy is strict. The proof of part (2) in [MT97b] is based on automata theory rather than Ehrenfeucht–Fraïssé games.

Schweikardt [Sch97] then improved part (1) by showing that a certain at least k-fold exponential function can be defined by a monadic Σ_k^{P} sentence and by a monadic Π_k^{P} sentence, with the corollary that the Boolean closure of monadic $\Sigma_{k-1}^{\mathrm{P}}$ is properly contained in the intersection of monadic Σ_k^{P} and monadic Π_k^{P}.

A result of Matz and Thomas that we use in this section is:

THEOREM 12.2 [MT97b]. *For all $k \geqslant 1$, there is a function that is at least k-fold exponential and that can be defined by a monadic $\Sigma_1^{TC(1)}$ sentence.*

Combining Theorem 12.2 with Proposition 10.2 gives:

THEOREM 12.3. *For all k, there is a function that is at least k-fold exponential and that can be defined by a closed monadic NP sentence.*

(This result was shown independently by Matz and Thomas [MT97a].) Theorems 12.1 and 12.3 reveal that the growth method cannot be used to prove strictness of the monadic closure hierarchy (or even prove that it does not collapse to the first level) in the same way that the growth method was used to prove strictness of the monadic hierarchy. This is because the growth method does not distinguish the closed monadic Σ_k^{P} classes for different k, as we now explain. Even closed monadic Σ_1^{P} sentences can define functions having at least k-fold exponential growth for arbitrarily large k (Theorem 12.3). And no matter how large k is, for every closed monadic Σ_k^{P} sentence defining a function f, there is an ℓ such that f has at most ℓ-fold exponential growth (this follows from Theorem 12.1(2)).

The new result in this section is that FO(MNP) sentences can define at most double-exponential growth.

THEOREM 12.4. *If the function f is defined by a FO(MNP) sentence, then f is at most 2-fold exponential.*

The proof of Theorem 12.4 is outlined below. An immediate corollary of Theorems 12.3 and 12.4 is:

COROLLARY 12.5. *There is a grid property that is in closed monadic NP but not in the first-order closure of monadic NP.*

This is the analog of Theorem 1.2 for grids. By encoding grids as undirected graphs, separations of certain descriptive complexity classes over grid properties can be translated to similar separations over undirected graph properties [MST]. Thus (leaving the details to the interested reader), it follows from Corollary 12.5 that the separation of closed monadic NP from the first-order closure of monadic NP holds also over undirected graph properties. Thus, the encoding method can be used to obtain Theorem 1.2 as a corollary of Corollary 12.5.

In a previous version of this paper, we left as open questions whether PFO(MNP) sentences or FO(MNP) sentences can define functions that are at least

2-fold exponential. Subsequently, these questions were answered by Matz [Mat98], who showed that a PFO(MNP) sentence (and, therefore, a FO(MNP) sentence) can define a function that is at least 2-fold exponential. He also gives an alternate (automata-theoretic rather than game-theoretic) proof of Theorem 12.4. By combining these results with Theorem 12.1(2), Theorem 12.3, and encoding of grids by undirected graphs, he obtains alternate proofs of Theorems 1.1 and 1.2. Matz [Mat98] also has results concerning the expressive power of the positive first-order closure of monadic Σ^{P}_{k} and the first-order closure of monadic Σ^{P}_{k}. Some of these new results extend Theorems 1.1 and 1.2.

In the remainder of this section we sketch the proof of Theorem 12.4. The proof is similar to (and simpler than) the proof of Theorem 9.6, and like our proof of Theorem 9.6, it uses Lemma 9.5. We view an (m, n)-grid as a chain of length n, where each point in the chain is actually a supernode consisting of a column of m points.

We first need a lemma giving a condition for the duplicator to have a winning strategy in a certain point-selecting game played on two chains. The winning condition for the duplicator is similar to the last three parts of the winning condition in the (r, t, M)-black/white game (the first part is not needed because in the new game the chains are not initially colored). Let A and B be two chains. Let t, M be positive integers. We say that two sequences $\langle a_{-1}, a_0, a_1, ..., a_k\rangle$ and $\langle b_{-1}, b_0, b_1, ..., b_k\rangle$ of points, where the points a_i are selected from A and the b_i are selected from B, have the *(t, M)-property* if:

1. $a_i = a_j$ iff $b_i = b_j$, for all i, j;

2. The sequences $\langle a_{-1}, a_0, a_1, ..., a_k\rangle$ and $\langle b_{-1}, b_0, b_1, ..., b_k\rangle$ appear in the same order in the two chains. That is, there is a permutation π: $\{-1, ..., k\} \rightarrow \{-1, ..., k\}$ such that, moving from left to right in the chain A (resp., B), the points occur in the order $a_{\pi(-1)}, ..., a_{\pi(k)}$ (resp., $b_{\pi(-1)}, ..., b_{\pi(k)}$); and

3. For $-1 \leqslant i < k$, let ℓ_i (resp., ℓ'_i) be the distance from $a_{\pi(i)}$ to $a_{\pi(i+1)}$ (resp., from $b_{\pi(i)}$ to $b_{\pi(i+1)}$). Then, for $-1 \leqslant i < k$, either (a) $\ell_i = \ell'_i$, or (b) $\ell_i > t$, and $\ell'_i > t$, and $\ell_i \equiv \ell'_i \bmod M$.

The rules of the (r, t, M)-chain game over chains A, B are as follows. The game has r rounds. In each round the spoiler selects a point in one chain and the duplicator responds by selecting a point in the other chain. Let a_i (resp., b_i) be the point selected during round i in A (resp., B). Let a_{-1} (resp., b_{-1}) be the left endpoint of A (resp., B), and let a_0 (resp., b_0) be the right endpoint of A (resp., B). The duplicator wins if $\langle a_{-1}, a_0, a_1, ..., a_r\rangle$ and $\langle b_{-1}, b_0, b_1, ..., b_r\rangle$ have the (t, M)-property.

LEMMA 12.6. *Let r, t, M be positive integers. If A, B are chains each of length greater than $2^{r+1}tM$ and whose lengths are congruent $\bmod\, tM$, then the duplicator has a winning strategy in the (r, t, M)-chain game over A, B.*

Proof. The proof is very similar to the proof of Lemma 9.4, and like that proof it uses Lemma 9.3. Let A be a chain of length n_A and B be a chain of length n_B, where $n_A, n_B > 2^{r+1}tM$ and $n_A \equiv n_B \bmod tM$. Let C be a chain of length tM. If

$n_A \not\equiv 0 \bmod tM$, let R be a chain of length j where $j \equiv n_A \bmod tM$ and $1 \leqslant j < tM$. Similar to the proof of Lemma 9.4, we view A as a chain A' of $\lceil n_A/tM \rceil$ supernodes, where each supernode corresponds to either C or R. If $n_A \equiv 0 \bmod tM$, then all supernodes of A' correspond to C; otherwise, the rightmost supernode of A' corresponds to R and the others to C. Define B' similarly in terms of B. Since $n_A \equiv n_B \bmod tM$, the rightmost supernodes of A' and B' are either both R or both C, and the other supernodes of both A' and B' are all C. Since $n_A, n_B > 2^{r+1}tM$, the chains A' and B' each have length greater than 2^{r+1}. So by Lemma 9.3, the duplicator has an order-respecting, endpoint-respecting winning strategy in the r-game over A', B'. This strategy gives the basis for the duplicator's winning strategy in the (r, t, M)-chain game over A, B, in a way identical to that in the proof of Lemma 9.4. That is, the duplicator uses the strategy of Lemma 9.3 to select supernodes, and selects points "within" supernodes by selecting corresponding points. The proof that this is a winning strategy in the (r, t, M)-chain game is identical to that given in Lemma 9.4. ∎

Proof sketch for Theorem 12.4. The goal is to show that for every FO(MNP) sentence φ there is a number z such that, for all m, n, if the (m, n)-grid satisfies φ and $n > s_2(zm)$, then there exists $n' > n$ such that the (m, n')-grid satisfies φ. It is immediate from this that φ does not define a function with growth exceeding $s_2(zm)$, since there is not a unique $n > s_2(zm)$ such that the (m, n)-grid satisfies φ. To obtain the goal it is enough (by Theorem 9.2) to show that for every r_1, c', r_2, there is z, such that for all m, n with $n > s_2(zm)$, there is $n' > n$, such that the duplicator has a winning strategy in the symmetric (r_1, c', r_2)-game over the (m, n)-grid and the (m, n')-grid.

We view an (m, n)-grid as a chain of length n where each point of the chain is actually a supernode consisting of a column of m points. Note that a coloring of the grid with c' colors corresponds to a supercoloring of the chain with $(c')^m$ colors.

Let r_1, c', r_2 be positive integers. For a given m, find t, M from Lemma 9.5 where the roles of c, r in the statement of Lemma 9.5 are played by $(c')^m, r_2$. Inspection of the proof of Lemma 9.5 shows that there is a number z', depending on c' and r_2 but not on m, such that $t, M \leqslant s_2(z'm)$ (intuitively, since each d-type in the proof of Lemma 9.5 is a vector of at most $2d-1$ colors, and since d depends only on $r = r_2$, the number N of d-types is bounded by a polynomial in the number $(c')^m$ of colors, and hence is at most exponential in m; by Stirling's formula, we then see that t and M are at most double-exponential in m). Choose the number z such that $2^{r_1+1}(s_2(z'm))^2 \leqslant s_2(zm)$ for all m, thus ensuring that $2^{r_1+1}tM \leqslant s_2(zm)$.

Let m, n be arbitrary with $n > s_2(zm)$, find t, M as described above, and note that $n > 2^{r_1+1}tM$. Let G_0 be the (m, n)-grid and G_1 be the $(m, n+tM)$-grid. The duplicator's strategy on G_0 and G_1 is as follows. During the first r_1 rounds, the duplicator uses the strategy of Lemma 12.6, where the roles of r, t, M are played by r_1, t, M. More precisely, whenever the spoiler selects a point x in one grid, the column of the duplicator's response y is determined by the column of x and the strategy of Lemma 12.6, and the row of y is the same as the row of x. The duplicator's coloring strategy and his point-selecting strategy during the last r_2 rounds are similar to those described in the proof of Theorem 9.6 (and the proof is simpler here since there is only one type of supernode, a column of m points, rather than two types). Let $i_1 < i_2 < \cdots < i_k$ be the indices of the columns of G_0 containing some point selected during the first r_1 rounds, including the indices 1 and n of the leftmost and rightmost columns even if no point in these two columns was selected. Define $i'_1 < i'_2 < \cdots < i'_{k'}$ similarly for G_1. Since the duplicator used a winning strategy in the (r_1, t, M)-chain game, it follows that $k = k'$, and whenever the spoiler selected a point in column i_j of G_0 the duplicator responded with a point in column i'_j of G_1, and vice versa. Assume $0 \leqslant j < k$, $\ell = i_{j+1} - i_j$, and $\ell' = i'_{j+1} - i'_j$. The duplicator's coloring and final point-selecting strategy on the subgrids consisting of columns in the interval $[i_j, i_{j+1})$ in G_0 and $[i'_j, i'_{j+1})$ in G_1 is: (i) the identity strategy if $\ell = \ell'$, or (ii) based on the strategy guaranteed by Lemma 9.5 if $\ell \neq \ell'$, since in this case $\ell > t$, $\ell' > t$, and $\ell \equiv \ell' \bmod M$. Further details are similar to those in the proof of Theorem 9.6 and are left to the reader.

13. CONCLUSIONS AND OPEN QUESTIONS

We have introduced a new descriptive complexity class, closed monadic NP. We argued on mathematical grounds that closed monadic NP is a natural monadic subclass of NP: what matters is the configuration of the higher-order quantifiers. In the case of closed monadic NP, what matters is that all second-order quantifiers are unary and existential, but it does not matter how first-order quantifiers are interleaved among the second-order quantifiers. Closed monadic NP is robust, in that it is closed under first-order quantification and existential unary second-order quantification, the two types of quantification that appear in the definition of monadic NP. We have demonstrated the expressive power of closed monadic NP by showing, for example, that it is a proper extension of the positive first-order closure of monadic NP (this follows from Theorem 1.2), and that it contains all properties that can be expressed by monadic $\Sigma_1^{TC(1)}$ sentences (Proposition 10.2). A major open issue is to find lower bounds involving closed monadic NP. Of course, one way to find such lower bounds, that is, show that a property is not in closed monadic NP, is to show that the property cannot be expressed in monadic second-order logic. For example, we noted in Section 3 that Turán [Tur84] has shown that the perfect matching property cannot be expressed in monadic second-order logic. While such a result reveals a limitation of unary second-order quantifiers, it does not reveal a limitation of existential unary second-order quantifiers when compared to general unary second-order quantifiers. Such a limitation would be revealed by a "yes" answer to the following question, which was noted previously in Section 10.

Open Question 1. Is there a property that can be expressed by a sentence of monadic second-order logic (equivalently, a property in the monadic hierarchy), but that does not belong to closed monadic NP?

As we noted in Section 10, the answer to this question is "yes" under the assumption that NP $\neq$ co-NP, since then non-3-colorability would be such a property, which is even in monadic co-NP. It is interesting to consider whether Open Question 1 can be resolved without making any complexity-theoretic assumptions.

Since closed monadic NP is a subclass of NP, the following refinement of Open Question 1 is also interesting.

Open Question 2. Is there a property that is in the intersection of NP and the monadic hierarchy, but that does not belong to closed monadic NP?

In other words, Open Question 2 asks whether closed monadic NP equals the intersection of NP and the monadic hierarchy. This is because closed monadic NP is contained in both NP and the monadic hierarchy, as we have already noted. As far as we know, the assumption NP $\neq$ co-NP is consistent with either answer to Open Question 2.

We have also defined the closed monadic hierarchy, a natural monadic version of the polynomial-time hierarchy, where we focus on the number of alternations of unary second-order quantifiers and allow arbitrary interleavings of first-order quantifiers among the unary second-order quantifiers. As noted in Section 11, it is an open question whether the closed monadic hierarchy is strict.

Open Question 3. Is it true that closed monadic $\Sigma^{\mathrm{P}}_{k+1}$ does not equal closed monadic Σ^{P}_{k} for all $k \geqslant 1$?

We showed (Theorem 11.3) that the answer to this question is "yes" under the complexity-theoretic assumption that the polynomial-time hierarchy is strict. Moreover, under the assumption that $\Sigma^{\mathrm{P}}_{k} \neq \Sigma^{\mathrm{P}}_{k+1}$, we showed that there is an undirected graph property that is in monadic $\Sigma^{\mathrm{P}}_{k+1}$ (and therefore in closed monadic $\Sigma^{\mathrm{P}}_{k+1}$) but not in closed monadic Σ^{P}_{k}. It would be interesting to settle this question without making any complexity-theoretic assumptions.

We noted in Section 12 that there are difficulties in using the growth method to establish a "yes" answer to any of the three questions above, because all closed monadic Σ^{P}_{k} classes have exactly the same power to define functions having arbitrarily large-fold exponential growth. It is still possible that looking at grid properties could be helpful in settling these questions, but the argument must consider details finer than the rough growth rate of the width-versus-height function.

To provide a "yes" answer to any of the questions above by using Ehrenfeucht–Fraïssé game arguments, it seems that we will have to deal with a new complication to the game. That is, we will have to deal with multiple coloring rounds, interleaved with pebbling rounds.

APPENDIX

Proof that k-round 3*-colorability is* Σ^{P}_{k}*-complete when restricted to graphs of maximum degree at most* $\max\{k, 4\}$. We describe a polynomial-time many-one reduction from a known Σ^{P}_{k}-complete problem to k-round 3-colorability with degree restriction $\max\{k, 4\}$. This problem, shown to be Σ^{P}_{k}-complete in [Sto77, Wra77], is the quantified Boolean formula problem restricted to k blocks of alternating quantifiers. To bring out the similarities with k-round 3-colorability, we state this problem as a game. An instance of *k-round QBF* is a propositional formula F in 3CNF form (conjunctive normal form with at most three literals per clause, where a literal is either a variable or the negation of a variable), together with a partition of the Boolean variables appearing in F into k sets $X_1, ..., X_k$. The question is whether player 1 has a winning strategy in the following k-round game: at round i for $1 \leqslant i \leqslant k$, player $i \bmod 2$ assigns a truth value to every variable in X_i; after all variables have been assigned truth values, the player who moved last (player $k \bmod 2$) wins the game if F is true under the chosen assignment.

The polynomial-time reduction from k-round QBF to k-round 3-colorability with degree restriction $\max\{k, 4\}$ is based on the reduction in [GJS76] from the 3CNF satisfiability problem to the 3-colorability problem with degree restriction 4. Let the 3CNF formula F and the variables-partition $X_1, ..., X_k$ be an instance of k-round QBF. To simplify the description of the reduction, it is done in several steps. The first step is to reduce k-round QBF to the modified k-round 3-coloring game described before the statement of Theorem 11.4. In the modified game, an instance consists of a graph G together with a partition of its vertices into k sets $V_1, ..., V_k$; vertices in V_i are colored by player $i \bmod 2$ during round i.

Let $x_1, ..., x_n$ be the variables appearing in F. The graph G contains vertices w and $x_i, \bar{x}_i, z_i$ for $1 \leqslant i \leqslant n$ (among other vertices). For $1 \leqslant i \leqslant n$, the graph contains the edges shown in Fig. 1. The vertex w is included in the set V_1 of vertices colored during round 1, so we can imagine that w is colored before any other vertex is colored. By convention, let 3 be the color of w. For a vertex v, say that v is "colored true" if it is colored either 1 or 2, and "colored false" if it is colored 3. The edges in Fig. 1 ensure that, in any legal 3-coloring of the vertices, x_i is colored true iff $\bar{x}_i$ is colored false. We want to construct the rest of G so that, given any "truth coloring" of the vertices $x_1, ..., x_n$, the coloring can be extended to a legal 3-coloring of G iff the corresponding truth assignment to the variables makes F true. For this purpose, we use a graph H from [GJS76] and shown in Fig. 2. This graph has the properties: (1) for any legal 3-coloring of H in which a_1, a_2, a_3 are all colored with the same color γ, the vertex y_6 must be colored γ; and (2) for any coloring of a_1, a_2, a_3 such that at least one of these vertices is colored γ, the coloring can be extended to a legal 3-coloring of H in which y_6 is colored γ. Let $\{a_{j1}, a_{j2}, a_{j3}\}$ be the literals in the jth clause of F for $1 \leqslant j \leqslant m$. (If a clause contains only one or two literals, then repeat one literal in the clause to produce a clause containing exactly three literals.) The graph G contains, for $1 \leqslant j \leqslant m$, a subgraph on vertices $\{y_{j1}, ..., y_{j6}\}$ isomorphic to the subgraph induced by $\{y_1, ..., y_6\}$ in Fig. 2. In addition, a_{jp} is connected to y_{jp} for $p = 1, 2, 3$, and y_{j6} is connected to w. Given an assignment of colors to the vertices $w, x_1, ..., x_n$, the properties of H imply that this coloring can be extended to a legal 3-coloring of all of G iff the "truth coloring" of the vertices $x_1, ..., x_n$ corresponds to an assignment of truth values to the variables

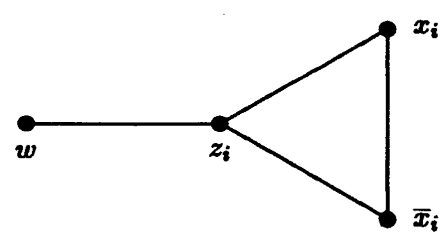

FIG. 1. The part of G representing the variables of F.

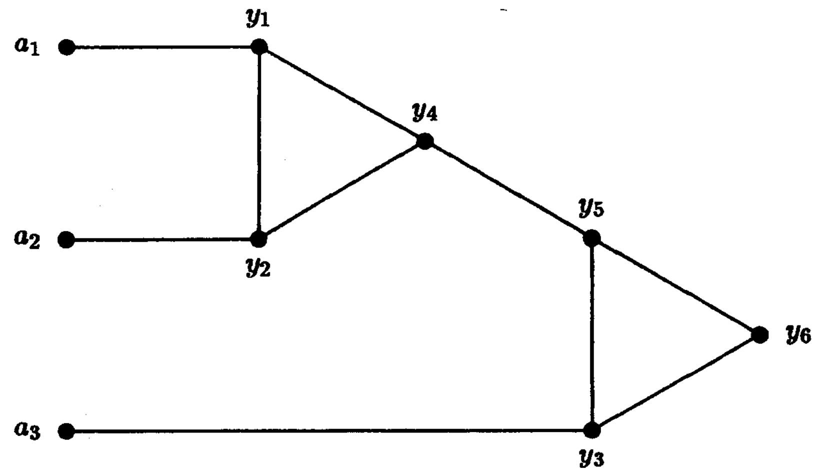

FIG. 2. The graph H.

that makes F true. Recall that the color of w is 3 (false) by convention. For example, if all the literals in the jth clause are colored 3 (false), then y_{j6} must be colored 3, but this cannot be a legal 3-coloring because y_{j6} is connected to w. The partition of vertices is obtained from the partition of variables as follows: $V_1 = X_1 \cup \{w\}$; $V_i = X_i$ for $1 < i < k$; and V_k contains all remaining vertices of G. (The vertex w is colored in the first round to fix a correspondence between truth values and colors for this and future rounds.) Player $k \bmod 2$ (the player moving last) wins the k-round QBF game on instance $(F, X_1, ..., X_k)$ iff player $k \bmod 2$ wins the modified k-round 3-coloring game on instance $(G, V_1, ..., V_k)$, since in the former game player $k \bmod 2$ wants to make F true and in the latter game player $k \bmod 2$ wants to legally 3-color G.

The next step is to transform $(G, V_1, ..., V_k)$ to an instance $(G', V'_1, ..., V'_k)$ of the modified k-round 3-coloring game, where G' has maximum degree at most 4. Here we use a graph from [GJS76]. For every integer $d \geqslant 3$, a "vertex substitute" graph H_d is described in [GJS76] having the following properties: (1) H_d has $O(d)$ vertices, maximum degree 4 and minimum degree 2; (2) H_d has d special vertices called *outlets*, and each outlet has degree 2; and (3) H_d is 3-colorable, and in any legal 3-coloring of H_d all outlets must be colored the same. The graph G is transformed to G' by replacing each vertex u of degree $d \geqslant 4$ with the graph $H_d(u)$, a copy of H_d; the d edges incident to u in G are incident to the outlets of $H_d(u)$ in G', with each outlet receiving one edge. The new vertex partition is obtained as follows: for $1 \leqslant i < k$, for every $u \in V_i$, if the degree of u is less than 4 (so u was not replaced by a copy of H_d) then $u \in V'_i$, or if the degree of u is $d \geqslant 4$ (so u was replaced by the graph $H_d(u)$), then V'_i contains some single outlet of $H_d(u)$; and V'_k contains all remaining vertices of G'. Note that, since every outlet of H_d has degree 2, every vertex in V'_i for $1 \leqslant i < k$ has degree at most 3.

The next step is to transform $(G', V'_1, ..., V'_k)$ to a graph G'', an instance of the original k-round 3-coloring game, where the degree of each vertex determines the round during which it is colored. As part of this transformation, the degrees of certain vertices must be increased. To increase the degree of vertex u from d to d', add $d' - d$ new edges incident to u. For each added edge, the other endpoint of the edge is connected to an arbitrary vertex of a copy of $K_{k,k}$, the complete bipartite graph with k vertices in each part of the bipartition. A separate copy of $K_{k,k}$ is used for each added edge. Since each vertex of $K_{k,k}$ has degree k, all vertices of the added $K_{k,k}$'s will be colored during the last round, round k. Since $K_{k,k}$ is 2-colorable, player $k \bmod 2$ will always be able to legally color all the $K_{k,k}$'s, regardless of how the rest of the graph is colored. If $k = 1$, then $G'' = G'$ (since the 1-round game is equivalent to 3-colorability). Assume $k \geqslant 2$.

Vertices in V_i' for $i \geqslant 3$ are easy to handle. Recall that all vertices in V_i' have degree at most 3 for $1 \leqslant i < k$, and all vertices in V_k' have degree at most 4. For $3 \leqslant i \leqslant k$ and each $u \in V_i'$, if the degree of u is less than i then the degree of u is increased to i. After this is done, each vertex in V_i' has degree exactly i for $3 \leqslant i < k$, each vertex in V_k' has degree exactly k if $k \geqslant 4$, and each vertex in V_k' has degree at least k and at most 4 if $k \leqslant 3$.

Vertices $u \in V_1' \cup V_2'$ require some different gadgets. Figure 3 shows how such vertices are handled in the two cases k even and k odd. In this figure, a solid square indicates a vertex whose degree has been increased (if needed) so that its degree is at least k, so all square vertices are colored during the last round. All edges and vertices other than u in this figure are added to G' to obtain G''.

Consider first the case k even, so player 2 colors during the last round, and player 2 wants the final coloring to be legal. Consider $u \in V_1'$. If player 1 has a winning strategy in the game on G' in which u is colored γ during round 1, then in

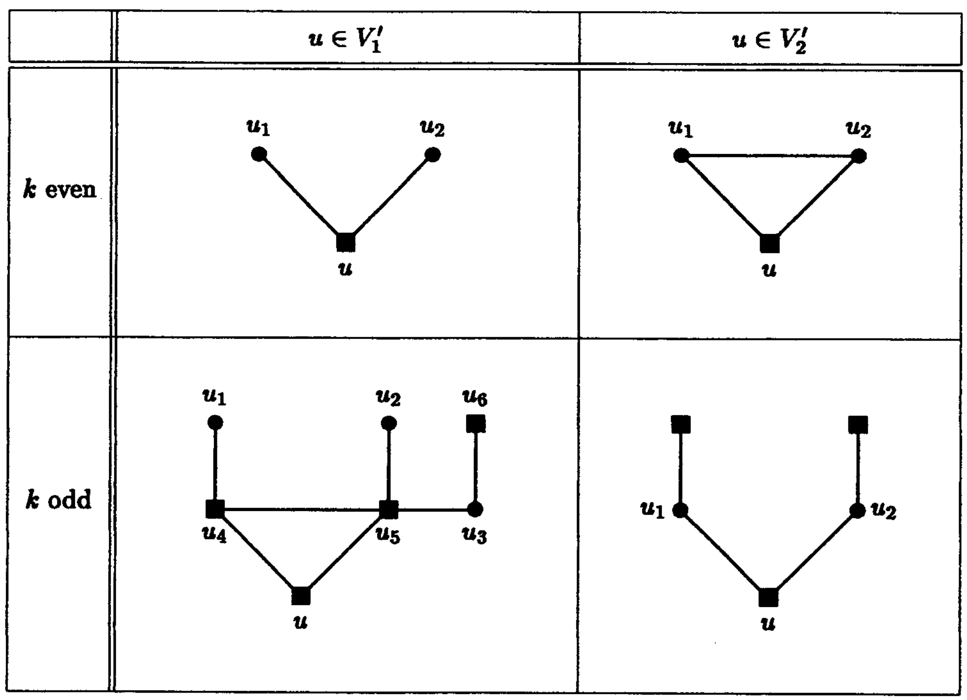

FIG. 3. The transformation from G' to G'' for $u \in V_1' \cup V_2'$. Squares indicate vertices of degree at least k.

the game on G'' player 1 colors u_1 and u_2 with the two colors other than γ. (u_1 and u_2 are colored during round 1 because they have degree 1 in this case.) This forces player 2 to color u with color γ during the last round. If player 1 does not have a winning strategy, then he cannot gain an advantage by coloring u_1 and u_2 the same, since this just gives more freedom to player 2 when he colors u. Now consider $u \in V'_2$. Since player 2 wants the final coloring to be legal, he must color u_1 and u_2 differently, thus forcing the color of u. (If $k = 2$, the transformation for $u \in V'_2$ is not needed.)

Now consider the case k odd, so player 1 colors during the last round, and player 1 wants the final coloring to be legal. Consider $u \in V'_1$. If player 1 has a winning strategy in the game on G' in which u is colored γ, then in the game on G'' player 1 colors u_1 and u_2 with γ. This forces player 1 to color u with γ during the last round, and u, u_4, u_5, u_6 can all be legally colored regardless of how player 2 colors u_3 during round 2. Suppose that player 1 does not have a winning strategy on G'. If u_1 and u_2 are colored the same, then again the color of u is forced. If u_1 and u_2 are not colored the same, then during round 2 player 2 colors u_3 with the same color as u_1. This forces u to have the same color as u_1 in any legal 3-coloring. Since player 2 has a winning strategy, he does not care how u is colored. The situation this construction avoids is allowing player 1 to delay choosing the color of u until after round 1. The argument for $u \in V'_2$ in this case, k odd, is identical to the argument for $u \in V'_1$ in the previous case, k even. This completes the transformation from k-round QBF to k-round 3-colorability.

It remains to prove the second part of Theorem 11.4 by showing that degree $\max\{k, 4\}$ suffices. The graph G'' in the above construction has maximum degree at most $\max\{k+1, 5\}$. We first transform G'' to a graph $G^{(3)}$ having maximum degree at most $\max\{k, 5\}$. Each copy of $K_{k,k}$ that was added to increase degrees contains one vertex of degree $k+1$ (the vertex a at which the copy of $K_{k,k}$ is attached to the rest of the graph), and these are the only vertices of degree $k+1$ if $k \geqslant 5$. By removing from each copy of $K_{k,k}$ one edge that is incident on its attachment vertex a, the maximum degree becomes at most $\max\{k, 5\}$. This reduces the degree of one vertex in each $K_{k,k}$ to $k-1$ (a vertex in the part of the bipartition not containing a) thus allowing player $(k-1) \bmod 2$ to choose the color of this vertex. It is easy to see, however, that player $k \bmod 2$ can still legally color all the $K_{k,k}$'s.

The final step is to transform $G^{(3)}$ (having maximum degree at most $\max\{k, 5\}$) to a graph $G^{(4)}$ having maximum degree at most $\max\{k, 4\}$. That is, for $k \leqslant 4$ we must reduce the degree from 5 to 4. We again use the "vertex substitute" graph H_d described in a previous step of the reduction. Each vertex u of degree 5 in $G^{(3)}$ is replaced by $H_5(u)$, a copy of H_5, and the edges incident to u in $G^{(3)}$ are incident to the outlets of $H_5(u)$ in $G^{(4)}$, with each outlet receiving one edge. Since $k \leqslant 4$, each such vertex u (having degree 5 in $G^{(3)}$) is colored during the last round when the game is played on $G^{(3)}$. To ensure that all vertices of $H_5(u)$ are colored during the last round when the game is played on $G^{(4)}$, if v is a vertex of $H_5(u)$ where the degree of v is less than 4, the degree of v is increased to 4 using methods already described, so that the resulting graph has degree at most 4.

ACKNOWLEDGMENTS

We are grateful to Prabhakar Raghavan, Thomas Schwentick, and Moshe Vardi for helpful comments.

REFERENCES

[AF90] M. Ajtai and R. Fagin, Reachability is harder for directed than for undirected finite graphs, *J. Symbolic Logic* **55**, No. 1 (1990), 113–150.

[AF97] S. Arora and R. Fagin, On winning strategies in Ehrenfeucht–Fraïssé games, *Theoret. Comput. Sci.* **174** (1997), 97–121.

[BH91] S. R. Buss and L. Hay, On truth-table reducibility to SAT, *Inform. and Comput.* **91** (1991), 86–102.

[CGH+88] J. Y. Cai, T. Gundermann, J. Hartmanis, L. A. Hemachandra, V. Sewelson, K. Wagner, and G. Wechsung, The Boolean hierarchy I: Structural properties, *SIAM J. Comput.* **17** (1988), 1232–1252.

[Cos93] S. S. Cosmadakis, Logical reducibility and monadic NP, *in* "Proc. 33rd IEEE Symp. on Foundations of Computer Science, 1993," pp. 52–61.

[Cou94] B. Courcelle, The monadic second order logic of graphs VI: On several representations of graphs by relational structures, *Discrete Appl. Math.* **54** (1994), 117–149.

[DLS98] A. Durand, C. Lautemann, and T. Schwentick, Subclasses of binary NP, *J. Logic Comput.* **8** (1998), 189–207.

[dR87] Michel de Rougemont, Second-order and inductive definability on finite structures, *Z. Math. Logik Grundl. Math.* **33** (1987), 47–63.

[Ehr61] A. Ehrenfeucht, An application of games to the completeness problem for formalized theories, *Fund. Math.* **49** (1961), 129–141.

[End72] H. B. Enderton, "A Mathematical Introduction to Logic," Academic Press, New York, 1972.

[Fag74] R. Fagin, Generalized first-order spectra and polynomial-time recognizable sets, *in* "Complexity of Computation, SIAM-AMS Proceedings" (R. M. Karp, Ed.), Vol. 7, pp. 43–73, 1974.

[Fag75] R. Fagin, Monadic generalized spectra, *Z. Math. Logik Grundl. Math.* **21** (1975), 89–96.

[Fag94] R. Fagin, Problem proposed at Oberwolfach Conference on Finite Model Theory, 1994.

[Fag97] R. Fagin, Comparing the power of games on graphs, *Math. Logic Quart.* **43** (1997), 431–455.

[Fra54] R. Fraïssé, Sur quelques classifications des systèmes de relations, *Publ. Sci. Univ. Alger. Sér. A* **1** (1954), 35–182.

[FSV95] R. Fagin, L. Stockmeyer, and M. Y. Vardi, On monadic NP vs. monadic co-NP, *Inform. and Comput.* **120** (1995), 78–92.

[GJS76] M. Garey, D. S. Johnson, and L. J. Stockmeyer, Some simplified NP-complete graph problems, *Theoret. Comput. Sci.* **1** (1976), 237–267.

[Gol80] M. Golumbic, "Algorithmic Graph Theory and Perfect Graphs," Academic Press, New York, 1980.

[Han65] W. Hanf, Model-theoretic methods in the study of elementary logic, *in* "The Theory of Models" (J. Addison, L. Henkin, and A. Tarski, Eds.), pp. 132–145, North Holland, Amsterdam, 1965.

[Imm89] N. Immerman, Descriptive and computational complexity, *in* "Computational Complexity Theory, Proc. Symp. Applied Math." (J. Hartmanis, Ed.), Vol. 38, pp. 75–91, Amer. Math. Soc., 1989.

[Mar99] J. Marcinkowsky, Directed reachability: From Ajtai–Fagin to Ehrenfeucht–Fraïssé games, *in* "Proceedings of the Annual Conference of the European Association for Computer Science Logic (CSL99)," 1999.

[Mat98] O. Matz, "First-Order Closure and the Monadic Second-Order Alternation Hierarchy," Technical Report 9807, Institut für Informatik und Praktische Mathematik, Christian-Albrechts-Universität Kiel, 1998.

[MP96] J. A. Makowsky and Y. B. Pnueli, Arity vs alternation in second order logic, *Ann. Pure and Appl. Logic* **78** (1996), 189–202.

[MR84] R. H. Möhring and F. J. Radermacher, Substitution decomposition for discrete structures and connections with combinatorial optimization, *Ann. Discrete Math.* **19** (1984), 257–356.

[MST] O. Matz, N. Schweikardt, and W. Thomas, The monadic quantifier alternation hierarchy over grids and graphs, *Inform. and Comput.*, to appear.
Note on the number of monadic quantifiers in monadic Σ^1_1, *Informa. Process. Lett.* **53** (1995), 337–339.

[MT97a] O. Matz and W. Thomas, Personal communication, 1997.

[MT97b] O. Matz and W. Thomas, The monadic quantifier alternation hierarchy over graphs is infinite, *in* "Proc. 12th IEEE Symposium on Logic in Computer Science," pp. 236–244, 1997.

[Ott95] M. Otto, Note on the number of monadic quantifiers in monadic Σ^1_1, *Informa. Process. Lett.* **53** (1995), 337–339.

[Sch95] T. Schwentick, Graph connectivity, monadic NP and built-in relations of moderate degree, *in* "Proc. 22nd International Colloq. on Automata, Languages, and Programming," pp. 405–416, 1995.

[Sch96] T. Schwentick, On winning Ehrenfeucht games and monadic NP, *Annals Pure Appl. Logic* **79** (1996), 61–92.

[Sch97] N. Schweikardt, The monadic quantifier alternation hierarchy over grids and pictures, *in* "Proceedings of the Annual Conference of the European Association for Computer Science Logic (CSL97)," Lecture Notes in Computer Science, Vol. 1414, pp. 441–460, Springer-Verlag, Berlin, 1997.

[Sho67] J. R. Shoenfield, "Mathematical Logic," Addison–Wesley, Reading, MA, 1967.

[Spi92] J. Spinrad, P_4-trees and substitution decomposition, *Discrete Appl. Math.* **39** (1992), 263–291.

[Sto77] L. J. Stockmeyer, The polynomial-time hierarchy, *Theoret. Comput. Sci.* **3** (1977), 1–22.

[Tur84] G. Turán, On the definability of properties of finite graphs, *Discrete Math.* **49** (1984), 291–302.

[Wag90] K. W. Wagner, Bounded query classes, *SIAM J. Comput.* **19** (1990), 833–846.

[Wra77] C. Wrathall, Complete sets and the polynomial-time hierarchy, *Theoret. Comp. Sci.* **3** (1977), 23–33.

Printed by Catherine Press, Ltd., Tempelhof 41, B-8000 Brugge, Belgium