DIMACS Series in Discrete Mathematics
and Theoretical Computer Science
Volume 31, 1997

# Easier Ways to Win Logical Games

Ronald Fagin
IBM Almaden Research Center
650 Harry Road
San Jose, California 95120-6099
email: fagin@almaden.ibm.com

fagin@almaden.ibm.com

URL: http://www.almaden.ibm.com/cs/people/fagin/

**Abstract**

The key tool in proving inexpressibility results in finite-model theory is Ehrenfeucht-Fraïssé games. This paper surveys various game-theoretic techniques and tools that lead to simpler proofs of inexpressibility results. The focus is on first-order logic and monadic NP.

## 1 Introduction

The *computational complexity* of a problem is the amount of resources, such as time or space, required by a machine that solves the problem. Complexity theory traditionally has focused on the computational complexity of problems. A more recent branch of complexity theory focuses on the *descriptive complexity* of problems, which is the complexity of describing problems in some logical formalism [Imm89]. One of the exciting developments in complexity theory is the discovery of a very intimate connection between computational and descriptive complexity. In particular, the author showed [Fag74] that the complexity class NP coincides with the class of properties of finite structures expressible in existential second-order logic, otherwise known as $\Sigma_1^1$. Because of this connection, a potential method of proving lower bounds in complexity theory is to prove inexpressibility results in the corresponding logic.

This issue of *expressive power* is fundamental in mathematical logic: given a class $\mathcal{C}$ of sentences and a class $\mathcal{M}$ of structures, we wish to know which properties $\mathcal{S}$ can be expressed by the sentences in $\mathcal{C}$ about the structures in $\mathcal{M}$. For example, let $\mathcal{C}$ be the class of all first-order sentences, let $\mathcal{M}$ be the class of all finite graphs, and let the property $\mathcal{S}$ be connectivity: then the question is whether there is a first-order sentence that is true about all finite graphs that are connected, but false about all finite graphs that are not connected (in this case, the answer is "No" [Fag75]).

We are interested in both positive results (which say that certain properties can be expressed) and negative results (which say that certain properties cannot be expressed). To prove a positive result, it is sufficient to exhibit a specific sentence in $\mathcal{C}$ and prove that this sentence expresses the property $\mathcal{S}$ over the given class $\mathcal{M}$ of structures. This is usually not very difficult. On the other hand, to prove a negative result, it is necessary to prove that there does not exist a sentence in $\mathcal{C}$ that expresses the property over $\mathcal{M}$. Since $\mathcal{C}$ is usually infinite, this means that the proof must simultaneously show that none of an infinite class of sentences "works". This is often a daunting task.

Fortunately, logicians have various tools in their arsenal to assist in proving inexpressibility results. These include the Compactness Theorem [End72], the Löwenheim-Skolem Theorem [End72], and Ehrenfeucht-Fraïssé games [Ehr61, Fra54]. If the class $\mathcal{M}$ consists only of finite structures (which is our main interest in this paper), then Ehrenfeucht-Fraïssé games are the only major tool available. (For some discussion on the failure of standard theorems in logic in the case of finite structures, see [Fag93, Gur84, Gur90].) The purpose of this paper is to discuss some results and techniques that assist in the use of Ehrenfeucht-Fraïssé games, and in particular that make the task of proving inexpressibility results easier.

There are several reasons why it is desirable to develop techniques that make the task of proving inexpressibility results easier. The first reason is to provide simpler proofs for inexpressibility results that are already known. This makes such results more understandable and accessible. The second reason is to make it possible to prove new and deeper inexpressibility results. Our hope is that we can develop such a powerful toolkit that we can eventually make a serious assault on such fundamental problems as the question of whether NP = co-NP. The development of new techniques can often accomplish both goals (of providing simpler proofs for known results, and of obtaining new results). For example, when an easier proof was given in [FSV95] for the result of [Fag75] that connectivity is not in monadic NP (which is defined shortly), this approach was used to show that the result remains true even in the presence of a larger class of built-in relations than was known before.

There are two classes $\mathcal{C}$ of sentences that we focus on here. The first consists of sentences in first-order logic: these are the primary sentences of interest in mathematical logic. The second class, which we shall discuss in more detail later, consists of sentences of the form $\exists A_1 \ldots \exists A_k \psi$, where each $A_i$ is a unary relation symbol and where $\psi$ is first-order. These are called *monadic* $\Sigma^1_1$ *sentences*, or *monadic NP* sentences [FSV95].

In Section 2, we give definitions and conventions. In Section 3, we discuss various sufficient conditions for the duplicator to have a winning strategy in a first-order Ehrenfeucht-Fraïssé game (played over two structures). These include Hanf's condition, as given by Fagin, Stockmeyer and Vardi [FSV95] (Section 3.1), Arora and Fagin's condition [AF94] (Section 3.2, and Schwentick's condition [Sch94] (Section 3.3). These three conditions are compared in Section 3.4. Roughly speaking, we could say that Hanf's condition requires isomorphic neighborhoods in the two structures; Arora and Fagin's condition requires approximately isomorphic neighborhoods in the two structures, along with other assumptions (such as that there be no small cycles); and Schwentick's condition requires that the structures be isomorphic, except in some small parts. In Section 4, we discuss techniques for proving inexpressibility results in monadic NP, such as Ajtai-Fagin games [AF90], where the rules are changed to help the duplicator. Some examples of the use of these techniques are given in Section 5. In Section 6, we make some additional comments. We give our conclusions in Section 7.

# 2 Definitions and conventions

A *language* $\mathcal{L}$ (sometimes called a *similarity type*, a *signature*, or a *vocabulary*) is a finite set $\{P_1, \ldots, P_p\}$ of relation symbols, each of which has an arity, along with a finite set $\{c_1, \ldots, c_z\}$ of constant symbols. An $\mathcal{L}$*-structure* (or *structure over* $\mathcal{L}$, or *structure of similarity type* $\mathcal{L}$, or simply *structure*) is a set $A$ (called the *universe*), along with a mapping associating a relation $R_i$ over $A$ with each $P_i \in \mathcal{L}$, where $R_i$ has the same arity as $P_i$, for $1 \leq i \leq p$, and associating a member of $A$ with each constant symbol $c_i \in \mathcal{L}$, for $1 \leq i \leq d$. We may call $R_i$ the *interpretation of* $P_i$ (and similarly for the constant symbols). If the point $a$ is the interpretation of the constant symbol $c_i$, then we may say that $a$ is *labeled* $c_i$. The structure is called *finite* if $A$ is.

We take a "graph" to be a structure where the language consists of a single binary relation symbol. Sometimes, such as in dealing with the reachability[1] problem, it is useful to take some liberties with standard terminology, by taking a "graph" to mean a directed graph with two distinguished points, labeled $s$ and $t$ respectively: then a graph is a structure where the language consists of a single binary relation symbol and two constant symbols, $s$ and $t$. We are also interested in "colored graphs", which are structures where the language includes also some finite number of unary relation symbols. If $G$ is a colored graph, where the interpretations of the unary relation symbols in the language are $U_1, \ldots, U_k$, then by the *color* of a point $a$ in the universe of $G$, we mean the set of $i$'s such that $a \in U_i$. Thus, there are $2^k$ colors.

Let $G$ be an $\mathcal{L}$-structure, and let $X$ be a subset of the universe of $G$. We write $G \restriction X$ for the substructure of $G$ induced by $X$. Thus, if $P_i$ is a relation symbol of $\mathcal{L}$, then the interpretation of $P_i$ in $G \restriction X$ is the set of tuples $t$ in $R_i$ such that every entry of $t$ is in $X$. In the case of constant symbols, there is a minor subtlety that deals with the situation where the vertex that is the interpretation of a constant symbol $c_j$ of $\mathcal{L}$ is not in the set $X$. So we take $G \restriction X$ to be an $\mathcal{L}'$-structure, where $\mathcal{L}'$ contains all of the relation symbols of $\mathcal{L}$, and only those constant symbols $c_j$ of $\mathcal{L}$ such that the interpretation of $c_j$ is in $X$. For such constant symbols $c_j$, the interpretation of $c_j$ in $G \restriction X$ is the same as the interpretation of $c_j$ in $G$.

We shall write $\langle x_1, \ldots, x_m \rangle$ to represent a tuple (an $m$-*tuple*) in an $m$-ary relation. In particular, we write $\langle x_1, x_2 \rangle$ to represent the directed edge from $x_1$ to $x_2$ in a directed graph. For an undirected graph, we write $(x_1, x_2)$ to represent the undirected edge between $x_1$ and $x_2$.

We use the usual Tarskian truth semantics (see, for example, [End72]) to define what it means for a structure $G$ to *obey* or *satisfy* a sentence $\sigma$, written $G \models \sigma$. It is assumed that $G$ and $\sigma$ are both over the same language $\mathcal{L}$. Let $\mathcal{M}$ be a fixed class of structures (such as the class of finite $\mathcal{L}$-structures for a given language $\mathcal{L}$). After this section, we shall suppress mention of $\mathcal{M}$. We define a *property* $\mathcal{S}$ to be a subset of $\mathcal{M}$ closed under isomorphism. For example, the property of connectivity is identified with the class of connected graphs. (In the

[1]The reachability (or "$(s,t)$-connectivity") problem is the problem of deciding, given a graph and two distinguished vertices $s$ and $t$ in it, whether there is a path from $s$ to $t$. In the case of directed graphs, the problem is called the *directed reachability problem*.

case of finite-model theory, where we restrict our attention to finite structures, "connectivity" would refer to the class of finite connected graphs.) We write $\overline{\mathcal{S}}$ for $\mathcal{M} \setminus \mathcal{S}$, the complement of $\mathcal{S}$ in $\mathcal{M}$. Let $\mathcal{C}$ be a class of sentences (such as the class of first-order sentences). When we say that the property $\mathcal{S}$ is *expressible* (or *definable*) in the class $\mathcal{C}$ of sentences over the class $\mathcal{M}$ of structures, we mean that there is a sentence $\sigma \in \mathcal{C}$ such that

1. if $\mathcal{A} \in \mathcal{S}$, then $\mathcal{A} \models \sigma$, and
2. if $\mathcal{A} \in \overline{\mathcal{S}}$, then $\mathcal{A} \not\models \sigma$.

# 3 First-order games

In this section, we focus on first-order Ehrenfeucht-Fraïssé games, and give three sufficient conditions for one player (the duplicator) to win. These conditions are based on techniques of Hanf [Han65] (and given a new interpretation by Fagin, Stockmeyer and Vardi [FSV95]), Arora and Fagin [AF94], and Schwentick [Sch94]. As we shall discuss, such techniques and conditions are valuable tools for obtaining inexpressibility results.

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. There are two *players*, called *the spoiler* and *the duplicator*, and two structures, $G_0$ and $G_1$. In the first round, the spoiler selects a point in one of the two structures, and the duplicator selects a point in the other structure. Let $p_1$ be the point selected in $G_0$, and let $q_1$ be the point selected in $G_1$. Then the second round begins, and again, the spoiler selects a point in one of the two structures, and the duplicator selects a point in the other structure. Let $p_2$ be the point selected in $G_0$, and let $q_2$ be the point selected in $G_1$. This continues for $r$ rounds. After $r$ rounds, $p_1, \ldots, p_r$ have been selected in $G_0$, and $q_1, \ldots, q_r$ have been selected in $G_1$. If the language contains constant symbols $c_1, \ldots, c_z$, then let $p_{r+i}$ denote the interpretation in $G_0$ of $c_i$, and let $q_{r+i}$ denote the interpretation in $G_1$ of $c_i$, for $1 \leq i \leq z$. The duplicator wins if the substructure of $G_0$ induced by $p_1, \ldots, p_{r+z}$ is isomorphic to the substructure of $G_1$ induced by $q_1, \ldots, q_{r+z}$, under the function that maps $p_i$ onto $q_i$ for $1 \leq i \leq r+z$. That is, the duplicator wins precisely if (a) $p_i = p_j$ iff

$q_i = q_j$, for each $i$ and $j$, and (b) $\langle p_{i_1}, \ldots, p_{i_\ell} \rangle$ is a tuple in a relation in $G_0$ iff $\langle q_{i_1}, \ldots, q_{i_\ell} \rangle$ is a tuple in the corresponding relation in $G_1$, for each choice of $i_1, \ldots, i_\ell$. 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 shows why these games are of interest.

**Theorem 3.1** [Ehr61, Fra54] *$\mathcal{S}$ is expressible in first-order logic iff there is $r$ such that whenever $G_0 \in \mathcal{S}$ and $G_1 \in \overline{\mathcal{S}}$, then the spoiler has a winning strategy in the $r$-game over $G_0, G_1$.*

In practice, we make use of what is essentially the contrapositive of Theorem 3.1, to prove that some property is *not* expressible in first-order logic. That is, we use the following theorem, which follows immediately from Theorem 3.1.

**Theorem 3.2** *$\mathcal{S}$ is not expressible in first-order logic iff for every $r$, there are $G_0 \in \mathcal{S}$ and $G_1 \in \overline{\mathcal{S}}$ such that $G_0 \sim_r G_1$.*

We see from Theorem 3.2 that to prove first-order inexpressibility, we would like tools for showing that the duplicator has a winning strategy in an $r$-game. As we shall see, such tools are also valuable as a step in proving inexpressibility in richer logics, such as monadic NP. We now discuss three sufficient conditions for the duplicator to have a winning strategy in an $r$-game, that is, for showing that $G_0 \sim_r G_1$ for two structures $G_0, G_1$.

## 3.1 Hanf's condition

Fagin, Stockmeyer and Vardi [FSV95] provide a simple but very useful sufficient condition for guaranteeing that $G_0 \sim_r G_1$ for two structures $G_0, G_1$. The proof is based on a technique of Hanf [Han65]. They used this condition as a part of a simple proof that connectivity is not in monadic NP (much simpler than the author's original proof [Fag75]).

Let $G$ be an $\mathcal{L}$-structure, where $\mathcal{L}$ consists of the relation symbols $P_1, \ldots, P_p$, possibly along with some constant symbols $c_1, \ldots, c_z$, and where $R_i$ is the interpretation in $G$ of the relation symbol $P_i$, for $1 \leq i \leq p$. The *Gaifman graph* [Gai82] of $G$ is the undirected graph with the same universe as $G$, and with an edge $(x_1, x_2)$ whenever $x_1$ and $x_2$ are distinct and appear together in a tuple of some relation of $G$. Let $a$ and $b$ be two points in (the universe of) $G$. We say that $a$ and $b$ are *adjacent* (in $G$) if either $a = b$, or $(a, b)$ is an edge of the Gaifman graph of $G$. Intuitively, two points $a$ and $b$ are adjacent if they are either identical or directly related by some relation of $G$. The *degree* of a point in $G$ is defined to be the degree in the Gaifman graph of $G$.

Define $\mathit{Ball}(a, k)$, the *ball of radius $k$ about $a$*, recursively as follows:

$$\mathit{Ball}(a, 0) = \{a\}$$

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

Define the *$d$-type* of a point $a$ to be the isomorphism type of the ball of radius $d - 1$ about $a$ with $a$ as a distinguished point.[2] Thus, the points $a$ in $G_0$ and $b$ in $G_1$ have the same $d$-type precisely if

$$G_0 \restriction \mathit{Ball}(a, d-1) \cong G_1 \restriction \mathit{Ball}(b, d-1),$$

under an isomorphism mapping $a$ to $b$. We say that the structures $G_0$ and $G_1$ are *$d$-equivalent* if for every $d$-type $\tau$, they have exactly the same number of points with $d$-type $\tau$. Intuitively, $d$-equivalence corresponds to a type of local isomorphism.

The next theorem gives a useful sufficient condition ("Hanf's condition") for the duplicator to have a winning strategy in a first-order game.

**Theorem 3.3** [FSV95] *Let $r$ be a positive integer. There is a positive integer $d$ such that $G_0 \sim_r G_1$ whenever $G_0$ and $G_1$ are $d$-equivalent structures.*

In fact, as shown in [FSV95], we can take $d = 3^{r-1}$ in Theorem 3.3.[3]

[2]This means that we are effectively considering the open ball of radius $d$, rather than the closed ball of radius $d$. We are doing this for compatibility with [FSV95], where this choice was made for technical convenience.

[3]This assumes that, as in [FSV95], there are no constant symbols. If there are $z$ constant symbols, then we would take $d = 3^{r+z-1}$, since, intuitively, $z$ constant symbols effectively increase the number of rounds by $z$ (our definition of a winning strategy for the duplicator assumes effectively that there are "$z$ extra rounds" where the points that are interpretations of the $z$ constant symbols are selected). A similar comment applies to the estimates following Theorems 3.6, 3.7, and 3.8.

We now give a simple example of the use of Hanf's condition, to show that connectivity is not first-order. We remark that this example is sufficiently simple that each of the three conditions (Hanf's, Arora-Fagin's, and Schwentick's) that we consider in this paper can prove this result. We make use here of Hanf's condition. By Theorem 3.2, we need only show that for each $r$, there is a graph $G_0$ that is connected and a graph $G_1$ that is not connected, such that $G_0 \sim_r G_1$. Given $r$, find $d$ as in Theorem 3.3. Let $G_0$ be a cycle with $4d$ nodes, and let $G_1$ be the disjoint union of two cycles, each with $2d$ nodes. It is easy to see that every point in $G_0$ and $G_1$ has the same $d$-type. Since $G_0$ and $G_1$ have the same number of points, and all with the same $d$-type, it follows that $G_0$ and $G_1$ are $d$-equivalent. By Theorem 3.3 and our choice of $d$, it follows that $G_0 \sim_r G_1$, which was to be shown. Later, in Section 5, we shall see a variation (from [FSV95]) of this proof applied to colored cycles, as a part of a simple proof that connectivity is not in monadic NP.

For the sake of later comparisons with Arora and Fagin's approach, we now give a slightly different version of Theorem 3.3.

**Theorem 3.4** *Let $r$ be a positive integer. There is a positive integer $d$ such that $G_0 \sim_r G_1$ whenever $G_0$ and $G_1$ are structures of the same similarity type that have the same set of vertices and that satisfy the condition that each vertex has the same $d$-type in $G_0$ as in $G_1$.*

On the face of it, the assumptions of Theorem 3.4 are stronger than those of Theorem 3.3, since the assumptions of Theorem 3.4 demand that each vertex have the same $d$-type in $G_0$ as in $G_1$, rather than simply requiring that for each $d$-type $\tau$, the structures $G_0$ and $G_1$ have the same number of vertices of $d$-type $\tau$. However, we can apply Theorem 3.4 whenever we could apply Theorem 3.3, by simply replacing $G_1$ by an isomorphic copy of $G_1$ (with the same set of vertices as $G_0$, and with each vertex having the same $d$-type in $G_0$ as in $G_1$).

The condition in Theorem 3.3 (that the two structures are $d$-equivalent for some large $d$) is sufficiently strong that it can be used to obtain indistinguishability results not just in first-order logic, but in stronger logics, as we now discuss. Let us write $G_0 \approx_r G_1$ if it happens that not only does the duplicator have a winning strategy in the $r$-game over $G_0, G_1$, but also the duplicator's strategy in each round is bijective. This means that for each $i$ (with $0 \leq i \leq r-1$) and each choice of $p_1, q_1, \ldots, p_i, q_i$ (where, intuitively, $p_j$ is the point chosen in

$G_0$ in round $j$, and $q_j$ is the point chosen in $G_1$ in round $j$), there is a bijection $f$ from the universe of $G_0$ to the universe of $G_1$, such that

1. if the spoiler selects $p_{i+1}$ in $G_0$ in round $i+1$, then the duplicator selects $f(p_{i+1})$ in $G_1$ in round $i+1$, and
2. if the spoiler selects $q_{i+1}$ in $G_1$ in round $i+1$, then the duplicator selects $f^{-1}(q_{i+1})$ in $G_0$ in round $i+1$.

When the duplicator's strategy in each round is bijective, then we say that the duplicator has a bijective strategy. By a result of Hella [Hel92], the duplicator's having a bijective winning strategy is sufficient to imply inexpressibility in first-order logic extended by unary generalized quantifiers.[4] Thus, Hella's result is the following analogue to Theorem 3.2. Unlike Theorem 3.2, this theorem gives only a sufficient condition, not a necessary and sufficient condition.

**Theorem 3.5** [Hel92] *$\mathcal{S}$ is not expressible in first-order logic extended by unary generalized quantifiers if for every $r$, there are $G_0 \in \mathcal{S}$ and $G_1 \in \overline{\mathcal{S}}$ such that $G_0 \approx_r G_1$.*

We note that Hella defines the Ehrenfeucht-Fraïssé game for unary generalized quantifiers slightly differently from how we do. Rather than simply requiring that the duplicator have a bijective strategy, as we do, Hella requires the duplicator to exhibit the bijection in each round before the spoiler makes his move. Hella's requirement does not change the notion of $\approx_r$.

Immerman and Lander [IL90] defined a game, which we shall call the "counting game". The rules of the $r$-round counting game are as follows. On round $i$ (for $1 \leq i \leq r$), the spoiler selects a set of points in one structure, and the duplicator must respond with a set of points of the same cardinality in the other structure. Then the spoiler selects a point in the set chosen by the duplicator, and the duplicator selects a point in the set chosen by the spoiler. Let $p_i$ be the point selected in $G_0$, and let $q_i$ be the point selected in $G_1$. As before, the duplicator wins if the substructure of $G_0$ induced by $p_1, \ldots, p_r$ is isomorphic to the substructure of $G_1$ induced by $q_1, \ldots, q_r$, under the function that maps $p_i$ onto $q_i$ for $1 \leq i \leq r$.

[4] See [Hel92] for the definition of unary generalized quantifiers. With unary generalized quantifiers, it is possible to express sentences like "there are an even number of points $x$ such that $Px$ holds" and "the number of points $x$ where $Px$ holds is less than the number of points $y$ where $Qy$ holds".

It is clear that if the duplicator has a bijective winning strategy in the (first-order) $r$-game, then he also has a winning strategy in the $r$-round counting game: the duplicator responds to moves of the spoiler by using his bijection, in an obvious way. (For example, if $f$ is the bijection in round $i$ between the universe of $G_0$ and the universe of $G_1$, and if in round $i$ the spoiler selects the set $S$, which is a subset of the universe of $G_0$, then the duplicator responds by selecting $\{f(x) \mid x \in S\}$ as his chosen subset of the universe of $G_1$.) What is not so clear (but is true) is that the converse also holds.[5] Thus, the duplicator has a winning strategy in the $r$-round counting game over $G_0, G_1$ iff $G_0 \approx_r G_1$.

Nurmonen [Nur96] showed the following strengthening of Theorem 3.3.

**Theorem 3.6** [Nur96] *Let $r$ be a positive integer. There is a positive integer $d$ such that whenever $G_0$ and $G_1$ are $d$-equivalent structures, then $G_0 \approx_r G_1$.*

In fact, as shown in [Nur96], we can take $d = 3^r$ in Theorem 3.6.

Etessami [Ete95] considered the following problem of defining an order: in a structure with a successor relation and constant symbols $s$ and $t$, does $s$ precede $t$? He proved that order is not expressible in first-order logic extended by unary generalized quantifiers, in the presence of a built-in successor relation.[6] His proof proceeds as follows.

Define two graphs $G_0$ and $G_1$, each of which are long chains, where the $(i+1)$st point in the chain is the successor of the $i$th point in the chain. In the graph $G_0$, the point labeled $s$ precedes the point labeled $t$, and in $G_1$, the opposite is true. The graph $G_0$ has a long stretch of points, followed by $s$, followed by another long stretch, followed by $t$, followed by another long stretch; and similarly for $G_1$, with the roles of $s$ and $t$ reversed. Etessami makes use of the counting game, and gives a long, involved proof that the duplicator has a winning strategy. The inexpressibility result then follows.

[5] The fact that the duplicator has a bijective winning strategy in the $r$-game iff he has a winning strategy in the $r$-round counting game was apparently first noted by Hella (personal communication). It is not hard to see that the equivalence of the two games follows from Observation 5.3 in [CFI92]. The equivalence is noted in the full version of [Hel92], but only for the $\omega$-round version.

[6] Actually, Etessami stated his result by saying that order is not expressible in "first-order logic with counting" in the presence of a built-in successor relation, but his proof amounts to essentially the same thing.

Instead of Etessami's long proof, we now show that it is simple to use Theorem 3.6 to show that for each $r$, there is a pair $G_0, G_1$ as above such that $G_0 \approx_r G_1$. Etessami's result then follows immediately from Theorem 3.5. By Theorem 3.6, we need only show that for each $d$, there is a pair $G_0, G_1$ as above such that $G_0$ and $G_1$ are $d$-equivalent. Let $G_0$ have universe $\{1, \ldots, 3d-1\}$, where the interpretation of the successor relation is the usual successor relation restricted to $\{1, \ldots, 3d-1\}$, where the interpretation of $s$ is the point $d$, and where the interpretation of $t$ is the point $2d$. Thus, intuitively, $G_0$ consists of a chain, with $d-1$ points, followed by $s$, followed by $d-1$ more points, followed by $t$, followed by $d-1$ more points. We define $G_1$ the same, except that we reverse $s$ and $t$, so that $t$ precedes $s$. In each structure, a $d$-type consists of a sequence of points, where one of the points in the sequence may be labeled by $s$ or $t$. It is easy to see that for every $d$-type, $G_0$ and $G_1$ have exactly the same number of points with that $d$-type. That is, $G_0$ and $G_1$ are $d$-equivalent, which was to be shown.

Although Theorem 3.3 is sufficient for their purposes, Fagin, Stockmeyer and Vardi actually prove a slightly stronger version of this theorem. Instead of demanding that $G_0$ and $G_1$ be $d$-equivalent (that is, for each $d$-type $\tau$, have exactly the same number of points with $d$-type $\tau$), they show that it is sufficient instead to require only that for every $d$-type $\tau$, either $G_0$ and $G_1$ have the same number of points with $d$-type $\tau$, or else both have at least $m$ points with $d$-type $\tau$ (for some large $m$ that depends only on the number $r$ of rounds and the maximal degree of any point in $G_0$ and $G_1$). Intuitively, this latter condition says that for each $d$-type $\tau$, the structures $G_0$ and $G_1$ have the same number of points with $d$-type $\tau$, where we can count only as high as $m$. Thomas [Tho91] proves a similar result. Theorem 3.3 is also related to a result by Gaifman [Gai82], who proved that in a precise sense, first-order logic talks only about neighborhoods.

## 3.2 Arora and Fagin's condition

Arora and Fagin [AF94] introduced another sufficient condition for guaranteeing that $G_0 \sim_r G_1$ for two structures $G_0, G_1$. They used their condition as a part of a proof that directed reachability is not in monadic NP (much simpler than Ajtai and Fagin's original proof [AF90] of this result). Intuitively, their condition requires a weaker local isomorphism than does the Hanf condition, but at the expense of extra assumptions (such as that there are no small cycles).

Before we can state the theorem, we need to define the notions of a *cycle* in a structure and of the $(r, k)$*-color* of a vertex in a structure. In the case of (colored) graphs, a cycle is the usual notion of an undirected cycle, where we ignore the directions of the edges. The definition in the case of general structures is more complicated, and we defer it till later. The definition of the $(r, k)$-color is also simpler in the case of colored graphs, and we give it now; later we give the definition in the general case. We note that the case of colored graphs is what is needed when using Arora and Fagin's condition as a "subroutine" in proving that some graph property is not in monadic NP.

Let $r$ and $k$ be integers, and let $G$ be a colored graph. We now define the notion of the $(r, k)$-color of each vertex in $G$. Intuitively speaking, the $(r, k)$-color approximately describes a small neighborhood around the vertex. Define the $(r, 0)$-color of a vertex $v$ to be the color of $v$ in the colored graph, along with a description of whether or not there is an edge (a "self-loop") from the vertex $v$ to itself. If (as in the case of reachability) there are also distinguished vertices $s, t$, then the $(r, 0)$-color of a vertex $v$ also tells whether or not the vertex $v$ is the distinguished point labeled $s$, and whether or not the vertex $v$ is the distinguished point labeled $t$. Inductively, define the $(r, k+1)$-color of the vertex $v$ (where $k \geq 0$) to be (a) a description of its $(r, k)$-color, along with (b) a complete description, for each possible $(r, k)$-color $\tau$, as to whether there are $0, 1, \ldots, r-1$, or at least $r$ points $w$ with $(r, k)$-color $\tau$ such that $\langle v, w\rangle$ is an edge of the graph, but $\langle w, v\rangle$ is not an edge, (c) a complete description, for each possible $(r, k)$-color $\tau$, as to whether there are $0, 1, \ldots, r-1$, or at least $r$ points $w$ with $(r, k)$-color $\tau$ such that $\langle w, v\rangle$ is an edge but $\langle v, w\rangle$ is not an edge, and (d) a complete description, for each possible $(r, k)$-color $\tau$, as to whether there are $0, 1, \ldots, r-1$, or at least $r$ points $w$ with $(r, k)$-color $\tau$ such that $\langle v, w\rangle$ and $\langle w, v\rangle$ are each edges. Thus, the $(r, k+1)$-color of a vertex $v$ tells the $(r, k)$-color of $v$, and also tells how many vertices there are of each $(r, k)$-color with just an outedge from $v$, just an inedge into $v$, and both an outedge from and an inedge into $v$, where in all cases we do not count beyond $r$.

The $(r, k)$-color of a tuple $\langle x_1, \ldots, x_m\rangle$ is the tuple $\langle \tau_1, \ldots, \tau_m\rangle$, where $\tau_i$ is the $(r, k)$-color of $x_i$, for $1 \leq i \leq m$.

The next theorem gives Arora and Fagin's sufficient condition for the duplicator to have a winning strategy in the $r$-game over structures $G_0, G_1$.

**Theorem 3.7** [AF94] *Let $r, f$ be positive integers. There is a positive integer $k$ that depends only on $r$ such that $G_0 \sim_r G_1$ whenever $G_0$ and $G_1$ are structures of the same similarity type that have the same set of vertices and that satisfy the following conditions:*

1. *the degree of every vertex in $G_0$ or $G_1$ is at most $f$;*
2. *there is no cycle in either $G_0$ or $G_1$ of length less than $k$;*
3. *each vertex has the same $(r, k)$-color in $G_0$ as in $G_1$; and*
4. *if $e$ is a tuple that is present in some relation in one structure but not in the corresponding relation in the other structure, then there are at least $f^k$ tuples in both of these relations that have the same $(r, k)$-color as $e$.*

In fact, as shown in [AF94], we can take $k = 3^{2r}$ in Theorem 3.7.

A good example of where the Arora-Fagin condition might be applicable but the Hanf condition might not is when $G_0$ is a graph and $G_1$ is the result of deleting one edge of $G_0$ (this situation arises in the proof that directed reachability is not in monadic NP; see Example 2 of Section 5). In this example, the $d$-type in $G_0$ of each of the endpoints of the edge that is deleted to form $G_1$ would typically not be a $d$-type of any point of $G_1$.

Arora and Fagin also give a strengthening of Theorem 3.7, in which small cycles are allowed under certain circumstances. They use the strengthened version to deal with inexpressibility in the presence of certain built-in relations.

In the remainder of this subsection, we give definitions of "cycle" and "$(r, k)$-color" for structures $G$ that are not necessarily graphs. This (somewhat technical) material can be skipped by those interested only in the case of graphs.

If $t = \langle x_1, \ldots, x_k \rangle$ is a tuple, define $[t]$ to be the set $\{x_1, \ldots, x_k\}$ of points that appear in $t$. Define the *hypergraph associated with structure* $G$ to be a hypergraph $(V, F)$ whose universe $V$ is the same as the universe of $G$ and whose set $F$ of (hyper)edges is

$$\{[t] : t \text{ is a tuple in some relation of } G\}.$$

A *(simple) path of length $k$* between two points $u, v$ of $G$ consists of a set of edges $S_1, \ldots, S_k \in F$ and a set of points $x_1, \ldots, x_{k-1} \in V$ such that (i) the $x_i$'s are distinct from each other and from $u$ and $v$, (ii)

$u \in S_1$, (iii) $v \in S_k$, and (iv) $x_i \in S_i \cap S_{i+1}$, for $1 \leq i < k$. The *distance* between distinct points $u$ and $v$ is the smallest $k$ such that there is a path of length $k$ between them, and the distance between a point and itself is 0. In particular, an alternate way to define $Ball(a,k)$ in Section 3.1 is to take $Ball(a,k)$ to be all points whose distance from $a$ is at most $k$.

If $k \geq 3$, then a *cycle of length* $k$ in a structure $G$ is a path of length $k$ from a vertex to itself. (Shortly, we shall mention why cycles of length 1 or 2 are not considered.) Except for the fact that cycles of length 2 are not considered, this definition corresponds to Berge's notion [Ber76] of a cycle in a hypergraph. (There are various other notions of a cycle in a hypergraph that are not equivalent to Berge's; see [Fag83].) Note that if $G$ is a structure over a language with a single binary relation, then its hypergraph is an ordinary undirected graph, and the concept of distance and cycle are the familiar ones.

The notions of a cycle in structure $G$ and a cycle in the Gaifman graph of $G$ are different in general. For example, if there is a tuple $\langle x_1, x_2, x_3 \rangle$ in a ternary relation of a structure $G$ with all entries distinct, then there is a cycle of length 3 in the Gaifman graph (with edges $(x_1, x_2)$, $(x_2, x_3)$, and $(x_3, x_1)$), but not necessarily a cycle in $G$. In general, a cycle in a structure gives rise to a cycle in the Gaifman graph, but not vice versa. Note that an assumption of Theorem 3.7 is that there are *no* small cycles in the structure. Thus, the fact that the notion of "cycle" we have given is restrictive only increases the applicability of Theorem 3.7. This is also why cycles of length less than 3 are not considered; such very small cycles would have no effect on the theorem, and so we do not want to forbid them.

We now discuss how to define the $(r,k)$-color of each vertex $v$ in a structure $G$ over an arbitrary language $\mathcal{L}$. We begin with a preliminary notion. A *basic $m$-type* (among the $m$ variables $x_1, \ldots, x_m$) is a conjunction such that (a) for each $i$ and $j$ between 1 and $m$, exactly one of $x_i = x_j$ or $x_i \neq x_j$ is a conjunct, and (b) for each arity $\ell$, each relation symbol $P \in \mathcal{L}$ of arity $\ell$, and each choice of $i_1, \ldots, i_\ell$ where $1 \leq i_j \leq m$ for each $j$, exactly one of $Px_{i_1} \ldots x_{i_\ell}$ or $\neg Px_{x_1} \ldots x_{i_\ell}$ is a conjunct. Intuitively, a basic $m$-type tells exactly how the variables $x_1, \ldots, x_m$ relate to each other in a quantifier-free way. We say that the variable $x$ *has a positive occurrence* in the basic $m$-type $F$ if $Px_{i_1} \ldots x_{i_\ell}$ (as opposed to $\neg Px_{i_1} \ldots x_{i_\ell}$) is a conjunct of $F$ for some relation symbol $P \in \mathcal{L}$ and some variables $x_{i_1}, \ldots, x_{i_\ell}$ where $x \in \{x_{i_1}, \ldots, x_{i_\ell}\}$. We define a basic $m$-type of vertices (as opposed to variables) analogously. Specifically, if $v_1, \ldots, v_m$ are $m$ vertices of $G$, then we define their basic $m$-type to be a basic $m$-type among $m$ variables $x_1, \ldots, x_m$ that holds in $G$ when $x_1, \ldots, x_m$ are interpreted by $v_1, \ldots, v_m$ respectively. Intuitively, a basic $m$-type among the $m$ vertices $v_1, \ldots, v_m$ of $G$ tells exactly how these vertices relate to each other in $G$. Similarly, we define what it means for a vertex $v$ to have a positive occurrence in a basic $m$-type.

We are now ready to define the $(r,k)$-color of each vertex $v$ in a structure $G$ over an arbitrary language $\mathcal{L}$. Let $m$ be the largest arity among relation symbols of $\mathcal{L}$. The $(r,0)$-color of $v$ is a complete description of which relations of $G$ have the tuple $\langle v, \ldots, v\rangle$ as a member (where, of course, the length of the tuple is the arity of the relation), and which constant symbols label $v$. Inductively, define the $(r, k+1)$-color of the vertex $v$ (where $k \geq 0$) to be a description of its $(r,k)$-color, along with a complete description, for each possible choice $\tau_1, \ldots, \tau_{m-1}$ of $(r,k)$-colors and each possible basic $m$-type $F$ among $m$ vertices where $v$ has a positive occurrence, as to whether there are $0, 1, \ldots, r-1$, or at least $r$ choices of $\langle v_1, \ldots, v_{m-1}\rangle$ such that $v, v_1, \ldots, v_{m-1}$ have the basic $m$-type $F$ and $v_i$ has $(r,k)$-color $\tau_i$ for $1 \leq i \leq m-1$.

## 3.3 Schwentick's condition

Schwentick [Sch94] introduced still another sufficient condition for guaranteeing that $G_0 \sim_r G_1$ for two structures $G_0, G_1$. He used this condition and some variations of it to prove several results, most importantly the result that connectivity is not in monadic NP, even in the presence of a built-in linear order. The Hanf condition and the Arora-Fagin condition cannot be applied when there is a built-in linear order, since then all of the vertices are in a ball of radius one.

The intuition behind Schwentick's condition is as follows. Assume that $H_0 \sim_r H_1$. Assume further that we extend $H_0$ and $H_1$ "in the same way" to larger structures $G_0$ and $G_1$. Then $G_0 \sim_r G_1$. Before we present Schwentick's condition formally, we need some more notation and definitions.

Let $G$ be a structure. If $x, y$ are points of $G$, let $d_G(x,y)$ be the distance between $x$ and $y$ in $G$ (as defined in Section 3.2). For simplicity of notation, we shall write simply $d(x,y)$ for $d_G(x,y)$, since it should be clear which structure we are taking the distance in. A *substructure* $H$ of $G$ is one induced by a set of nodes (as in the definition of Ehrenfeucht-

Fraïssé games). If $x$ is a point of $G$, and $H$ is a substructure of $G$, define $d(x, H)$ to be the minimum of $d(x, y)$ over all points $y$ of $H$. Define $Ball(H, k)$ to be the set of points $x$ of $G$ such that they are within distance $k$ of $H$, that is, such that $d(x, H) \leq k$. Define $N(H, k)$ to be the substructure of $G$ induced by $Ball(H, k)$.

Let $G_0, G_1$ be structures with substructures $H_0, H_1$ respectively. Say that *the duplicator has a distance-respecting winning strategy in the $r$-game on $N(H_0, k)$ and $N(H_1, k)$* if he has a winning strategy in the $r$-game over $N(H_0, k), N(H_1, k)$ such that whenever $p_i$ (resp., $q_i$) is the point picked in round $i$ in $N(H_0, k)$ (resp., $N(H_1, k)$), then $d(p_i, H_0) = d(q_i, H_1)$. Say that *there is a distance-preserving isomorphism from $G_0 - H_0$ to $G_1 - H_1$* if there is an isomorphism $f$ between (a) the substructure of $G_0$ with universe those points not in $H_0$ and (b) the substructure of $G_1$ with universe those points not in $H_1$, such that $d(x, H_0) = d(f(x), H_1)$ for every point $x$ of $G_0 - H_0$.

We can now state Schwentick's sufficient condition.

**Theorem 3.8** [Sch94] *Let $r$ be a positive integer. There is a positive integer $k$ that depends only on $r$ such that $G_0 \sim_r G_1$ whenever $G_0$ and $G_1$ are structures of the same similarity type, with substructures $H_0, H_1$ respectively, that satisfy the following conditions:*

1. *the duplicator has a distance-respecting winning strategy in the $r$-game on $N(H_0, k)$ and $N(H_1, k)$; and*

2. *there is a distance-preserving isomorphism from $G_0 - H_0$ to $G_1 - H_1$.*

In fact, as shown in [Sch94], we can take $k = 2^r$ in Theorem 3.8.

As we noted earlier, Schwentick's condition (like Hanf's, and Arora-Fagin's) can be used to prove that connectivity is not first-order, by considering one cycle versus two cycles. We now show how, and note how the result can be strengthened. As before, by Theorem 3.2 we need only show that for each $r$, there is a graph $G_0$ that is connected and a graph $G_1$ that is not connected, such that $G_0 \sim_r G_1$. Given $r$, find $k$ as in Theorem 3.8. Let $G_0$ be a cycle with $4k+4$ nodes, and let $G_1$ be the disjoint union of two cycles, each with $2k+2$ nodes. Let $H_0$ be the subgraph of $G_0$ that contains two nodes that are as far apart as possible. Let $H_1$ be the subgraph of $G_1$ that contains two nodes that are in different cycles. It is easy to see that the conditions of Theorem 3.8 are satisfied. Hence $G_0 \sim_r G_1$, which was to be shown. In [Sch95], Schwentick shows how to extend this argument (again, by using only Theorem 3.8) to show that connectivity is not in monadic NP, even in the presence of built-in relations of degree $n^{o(1)}$. This strengthens Fagin, Stockmeyer and Vardi's result [FSV95] that connectivity is not in monadic NP, even in the presence of built-in relations of degree $(\log n)^{o(1)}$.

Schwentick presents several variations of his condition, that get increasingly powerful, but, unfortunately, also increasingly hard to understand. The most powerful (and hardest to understand) variation is used to prove his key result, that connectivity is not in monadic NP, even in the presence of a built-in linear order. At the end of this subsection, we comment on the idea behind how Theorem 3.8 must be modified to deal with a built-in linear order. Some modification is clearly required, since in the presence of a built-in linear order, each point has distance at most one from every other point. See [Sch95, Sch96] for other applications of Schwentick's method. Fortunately, the underlying proof of each of Schwentick's variations is essentially identical, and is elegant and easy to understand. Schwentick's argument is as follows.

At the beginning of the game we view the vertices in $H_0$ and $H_1$ as *inner vertices*, and the vertices outside of $Ball(H_0, 2^r)$ and $Ball(H_1, 2^r)$ as *outer vertices*. The other vertices are considered to be in a buffer area. The boundaries of the inner vertices and of the outer vertices may change on each round. Note that at the beginning of the game, the distance from every inner vertex to every outer vertex is greater than $2^r$.

What the duplicator does in a given round depends on whether the spoiler selects an inner vertex, an outer vertex, or a vertex in the buffer area. If the spoiler selects an inner vertex, then the duplicator responds, based on his winning strategy on the inner vertices. If the spoiler selects an outer vertex, then the duplicator responds, based on the isomorphism. If the spoiler selects a vertex in the buffer area, then there are two possibilities, depending on whether the vertex is closer to the inner vertices or the outer vertices. Assume without loss of generality that the spoiler selects vertex $p$ in $G_0$. Let $D$ be the distance from $p$ to the inner vertices, that is, the minimum of $d(p, x)$ where $x$ is an inner vertex of $G_0$.

Case 1: If $p$ is closer to the inner vertices than the outer vertices, then expand the inner vertices to include all of $Ball(H_0, D)$ and

*Ball*$(H_1, D)$. The duplicator now responds, based on his winning strategy on the (new) inner vertices.

Case 2: Otherwise, expand the outer vertices to include all vertices outside of *Ball*$(H_0, D-1)$ and *Ball*$(H_1, D-1)$. The duplicator now responds, based on the isomorphism.

It is straightforward to show, by induction on $i$, that after round $i$, the distance from every outer vertex to every inner vertex is more than $2^{r-i}$. In particular, at the end of the game, no inner vertex is adjacent to any outer vertex.

The duplicator wins the game, since (a) the substructure induced by vertices chosen during the game that are contained in the final set of inner vertices of $G_0$ is isomorphic to the analogous substructure in $G_1$; (b) the substructure induced by vertices chosen during the game that are contained in the final set of outer vertices of $G_0$ is isomorphic to the analogous substructure in $G_1$; and (c) there are no edges between the final set of inner vertices and the final set of outer vertices. This concludes the proof.

We close this subsection by giving the idea behind how Schwentick extends Theorem 3.8 to prove that connectivity is not in monadic NP, even in the presence of a built-in linear order. What we shall explain is not quite enough to prove this result, but at least it gives the proper intuition.[7] We take $G_0$ and $G_1$ in Theorem 3.8 to be structures that involve not only the graph relation but also the (built-in) linear ordering relation. The *distance* between two points is taken to be the distance using only the graph relation (and thus ignoring the linear ordering relation). This way, it is no longer the case that each point necessarily has distance at most one from every other point. A third condition is added to the two conditions of Theorem 3.8. This third condition is a homogeneity condition, which says that if $x$ and $y$ are points of $G_0$ such that the distance $d(x, H_0)$ is less than the distance $d(y, H_0)$, then $x$ is less than $y$ in the linear order (of course, we assume also that the symmetric condition holds for $G_1$). In particular, the points in $H_0$ form an initial segment of the linear order in $G_0$ (and similarly for $H_1$ in $G_1$). Once again, essentially the same proof shows that $G_0 \sim_r G_1$. The third condition ensures that in both graphs, after the game every inner vertex is less than every outer vertex in the linear order.

[7]The extension we now describe does not appear as such in any of Schwentick's papers. Schwentick mentioned this extension to the author in a private correspondence.

## 3.4 Comparison of the approaches

In this subsection, we discuss and compare the three sufficient conditions we have seen (Theorems 3.3, 3.7, and 3.8 and their variations) for the duplicator to have a winning strategy in a first-order game.

We begin by noting that the three conditions are incomparable: for each of the three conditions, there are situations where it can be applied but the other two cannot. Each corresponds to a different "reason" why the duplicator has a winning strategy.

Let us compare the Hanf condition (in the variation given by Theorem 3.4) with the Arora-Fagin condition (Theorem 3.7). We see that the assumptions of both theorems require that $G_0$ and $G_1$ have the same set of vertices. The Hanf condition requires isomorphic neighborhoods, whereas the Arora-Fagin condition requires only "approximately isomorphic neighborhoods" (by dealing only with $(r, k)$-colors, rather than isomorphism types). But the Arora-Fagin condition requires additional assumptions (such as that there be no small cycles). Just after the statement of Theorem 3.7, we mentioned an example where the Arora-Fagin condition is applicable but the Hanf condition is not.

Schwentick's condition requires even a stronger type of isomorphism than either the Hanf condition or the Arora-Fagin condition. Intuitively, Schwentick's condition requires that the structures be isomorphic, except in some small parts.

Historically, the importance of Schwentick's approach is that it is the first to be able to deal with a built-in linear order (this requires the strongest of Schwentick's variations). As we noted earlier, the Hanf condition and the Arora-Fagin condition cannot be applied in this case, since all of the vertices are in a ball of radius one.

We recommend that in applying Schwentick's approach, his underlying proof technique, as discussed in Section 3.3, be used rather than his theorems (except for the simplest variation, namely Theorem 3.8, and the unpublished extension mentioned at the end of Section 3.3). This is because, as we discussed, the stronger variations of his condition are hard to understand, whereas the proof technique is simple and elegant. By contrast, in applying Hanf's approach or Arora and Fagin's approach, we recommend that the theorems be used directly. This is because the proofs of correctness of these latter two approaches are much harder to understand and apply than the statements of the theorems.

# 4 Monadic NP

As mentioned earlier, the complexity class NP coincides with the class of properties of finite structures expressible in existential second-order logic [Fag74]. A consequence of this equivalence is that the famous question in complexity theory as to whether NP=co-NP is equivalent to the question in logic of whether existential and universal second-order logic have the same expressive power over finite structures, i.e., whether or not $\Sigma^1_1 = \Pi^1_1$. Since our best available tool for attacking the question of whether $\Sigma^1_1 = \Pi^1_1$ is Ehrenfeucht-Fraïssé games, Fagin, Stockmeyer and Vardi [FSV95] announced a program to build a toolkit of game-theoretic techniques. To help develop the toolkit, we restrict our attention to a "tractable" subclass of $\Sigma^1_1$, called *monadic NP*. We now give some definitions.

When we pass from first-order logic to second-order logic, we allow quantification over sets and relations. In particular, a $\Sigma^1_1$ *sentence* is a sentence of the form $\exists A_1 ... \exists A_k \psi$, where $\psi$ is first-order and where the $A_i$'s are relation symbols. As an example, we now construct a $\Sigma^1_1$ sentence that says that a graph (with edge relation denoted by $E$) is 3-colorable. In this sentence, the three colors are represented by the unary relation symbols $A_1$, $A_2$, and $A_3$. Let $\psi_1$ say "Each point has exactly one color". Thus, $\psi_1$ 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 $\psi_2$ say "No two points with the same color are connected by an edge". Thus, $\psi_2$ is

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

The $\Sigma^1_1$ sentence $\exists A_1 \exists A_2 \exists A_3 (\psi_1 \wedge \psi_2)$ then says "The graph is 3-colorable".

A $\Sigma^1_1$ sentence $\exists A_1 ... \exists A_k \psi$, where $\psi$ 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 class $\mathcal{S}$ of structures is said to be (*monadic*) $\Sigma^1_1$ if it is the class of all structures (of a given similarity type) that obey some fixed (monadic) $\Sigma^1_1$ sentence. When we restrict our attention to finite structures, a (monadic) $\Sigma_1^1$ class is also called a (monadic) generalized spectrum. Because of the equivalence between $\Sigma_1^1$ and NP, Fagin, Stockmeyer and Vardi [FSV95] refer to the collection of monadic $\Sigma_1^1$ classes (again, when we restrict attention to finite structures) as *monadic NP*. We often refer to a class of graphs by a defining property, for example, 3-colorability. As we saw above, 3-colorability is in monadic NP.

The author proved the first result about monadic NP, by showing that connectivity is not in monadic NP [Fag75]. We now discuss the original proof, in order to see the difficulties involved. Then we will see how the proof can be simplified by using various tools.

In [Fag75], the author introduced an Ehrenfeucht-Fraïssé game corresponding to 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 *$(c, r)$-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.

Note that 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 3.1.

**Theorem 4.1** [Fag75] *$\mathcal{S}$ is in monadic NP iff there are $c, r$ such that whenever $G_0 \in \mathcal{S}$ and $G_1 \in \overline{\mathcal{S}}$, then the spoiler has a winning strategy in the $(c, r)$-game over $G_0, G_1$.*

We now sketch how the author used Theorem 4.1 to prove that connectivity is not in monadic NP. Given $c, r$ it was shown that there are cycles $C_0, C_1$ such if $G_0 = C_0$ and $G_1 = C_0 \oplus C_1$ (where $\oplus$ represents the disjoint union) such that the duplicator can win the $(c, r)$-game over $G_0, G_1$. Since $G_0$ is connected and $G_1$ is not, this shows that connectivity is not in monadic NP.

The idea of the duplicator's coloring strategy was to color $C_0$ in $G_1$ by mimicking the coloring of $G_0$, and to color $C_1$ in a way where every

$d$-type in $C_1$ appears many times in $C_0$. The duplicator's pebbling strategy (that is, his strategy in the remaining $r$-game) was given explicitly. We note that for the duplicator's pebbling strategy, we could instead have made use of the extended version of Hanf's condition, mentioned at the end of Section 3.1.

For the sake of future discussions, let us consider what the difficulties are in this original proof. They are:

**D1:** The selection of the graphs $G_0, G_1$.[8]

**D2:** The duplicator's coloring strategy.

**D3:** The duplicator's pebbling strategy.

In addition to considering games over pairs $G_0, G_1$ of structures, Ajtai and Fagin [AF90] found it convenient, for reasons we shall see shortly, to consider games over a class $\mathcal{S}$. The rules of the game are as follows:

1. The duplicator selects a member of $\mathcal{S}$ to be $G_0$.
2. The duplicator selects a member of $\overline{\mathcal{S}}$ to be $G_1$.
3. The spoiler colors $G_0$ with the $c$ colors.
4. The duplicator colors $G_1$ with the $c$ colors.
5. The spoiler and duplicator play an $r$-game on the colored $G_0, G_1$.

We refer to this game as the *original* $(c, r)$ *game over* $\mathcal{S}$ (to contrast it with the Ajtai-Fagin $(c, r)$-game over $\mathcal{S}$, which we shall define shortly).

The next theorem follows easily from Theorem 4.1.

**Theorem 4.2** *$\mathcal{S}$ is in monadic NP iff there are $c, r$ such that the spoiler has a winning strategy in the original $(c, r)$-game over $\mathcal{S}$.*

[8] The issue is what the size of the cycles $C_0$ and $C_1$ should be. It is not sufficient that they simply be "sufficiently large". For example, we leave it to the reader to verify that if $C_0$ and $C_1$ are both odd cycles, if $G_0 = C_0$ and $G_1 = C_0 \oplus C_1$, and if $c \geq 2$ and $r \geq 3$, then the spoiler has a winning strategy in the $(c, r)$-game over $G_0, G_1$.

We now explain why Ajtai and Fagin allow $G_0$ and $G_1$ to be selected by the duplicator, rather than inputs to the game. By considering the choice of $G_0$ and $G_1$ to be moves of the duplicator, they were able to define a variation of $(c,r)$-games, in which the choice of $G_1$ by the duplicator is delayed until after the spoiler has colored $G_0$. They successfully used the new game to prove a result (that directed reachability is not in monadic NP) that they were not able to obtain using the original game. Their new game, called the *Ajtai-Fagin $(c,r)$-game*, is, on the face of it, easier for the duplicator to win. The rules of the new game are obtained from the rules of the $(c,r)$-game by reversing the order of two of the moves. Specifically, the rules of the Ajtai-Fagin $(c,r)$-game (over $\mathcal{S}$) are as follows.

1. The duplicator selects a member of $\mathcal{S}$ to be $G_0$.

2. The spoiler colors $G_0$ with the $c$ colors.

3. The duplicator selects a member of $\overline{\mathcal{S}}$ to be $G_1$.

4. The duplicator colors $G_1$ with the $c$ colors.

5. The spoiler and duplicator play an $r$-game on the colored $G_0, G_1$.

The winner is decided as before. Thus, in the Ajtai-Fagin $(c,r)$-game, the spoiler must commit himself to a coloring of $G_0$ with the $c$ colors before knowing what $G_1$ is. Intuitively, the Ajtai-Fagin game is harder for the spoiler to win than the original game. This can be made precise: it is shown in [Fag95] that in some situations, the spoiler requires strictly more resources (colors) to win the Ajtai-Fagin game than the original game. In spite of this, we have the following analogue to Theorem 4.2.

**Theorem 4.3** [AF90] *$\mathcal{S}$ 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 $\mathcal{S}$.*

Since the Ajtai-Fagin game is "easier for the duplicator to win", this makes our inexpressibility proofs easier. We now give two examples.

# 5 Examples

In this section, we discuss examples of the use of the techniques we have mentioned.

**Example 1: Connectivity.** Our first example deals with the proof that connectivity is not in monadic NP. We saw a sketch of the author's original proof in Section 4. We now give a simplified proof by Fagin, Stockmeyer and Vardi [FSV95].

**Simplified proof:** Let $\mathcal{S}$ be the class of connected graphs, and let $c, r$ be arbitrary. We now show that the duplicator has a winning strategy in the Ajtai-Fagin $(c, r)$-game over $\mathcal{S}$. It follows from Theorem 4.3 that $\mathcal{S}$ is not in monadic NP.

Let $d$ be given by Theorem 3.3 for this $r$. The duplicator chooses $G_0$ to be a directed cycle of length $n$, for a sufficiently large $n$. Let $\alpha_0, \alpha_1, \ldots, \alpha_{n-1}$ denote the points in order around the cycle, so that there is an edge from $\alpha_i$ to $\alpha_{i+1}$ for $0 \leq i < n$. Here and subsequently, subscripts are reduced modulo $n$ to belong to the interval $[0, n-1]$.

The spoiler now colors $G_0$ with $c$ colors. Let $\chi(\alpha_i)$ denote the color of $\alpha_i$. Assuming that $n \geq 2d$, the $d$-type of the point $\alpha_i$ in the resulting structure is fully described by the following vector of $2d - 1$ colors:

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

The number of possible $d$-types is some constant, depending on $c$ and $d$, but not on $n$. So it is clear that, for $n$ sufficiently large, there must be at least $4d$ points with the same $d$-type. Therefore, there must exist points $\alpha_p$ and $\alpha_q$ that have the same $d$-type and are at least distance $2d$ apart (that is, $\alpha_p \notin \mathit{Ball}(\alpha_q, 2d)$).

The duplicator now forms $G_1$, a pair of disjoint directed cycles, by pinching $G_0$ together at the points $\alpha_p$ and $\alpha_q$ (see Figure 1).

More precisely, let $G_1$ be a structure with universe consisting of $n$ distinct points $\beta_0, \beta_1, \ldots, \beta_{n-1}$. There is an edge from $\beta_i$ to $\beta_{i+1}$ for all $i$ with $0 \leq i < n$, $i \neq p$, and $i \neq q$, there is an edge from $\beta_p$ to $\beta_{q+1}$, and there is an edge from $\beta_q$ to $\beta_{p+1}$. There are no other edges. The duplicator's coloring of $G_1$ is given by $\chi(\beta_i) = \chi(\alpha_i)$ for all $i$.

Note that each component cycle of $G_0$ or $G_1$ contains at least $2d$ points, since $\alpha_p$ and $\alpha_q$ are at least distance $2d$ apart. Since also $\alpha_p$ and $\alpha_q$ have the same $d$-type, it follows that $\alpha_i$ and $\beta_i$ have the same $d$-type for all $i$, so $G_0$ and $G_1$ are $d$-equivalent. It follows from Theorem 3.3 that $G_0 \sim_r G_1$, so the duplicator wins. This concludes the proof. ∎

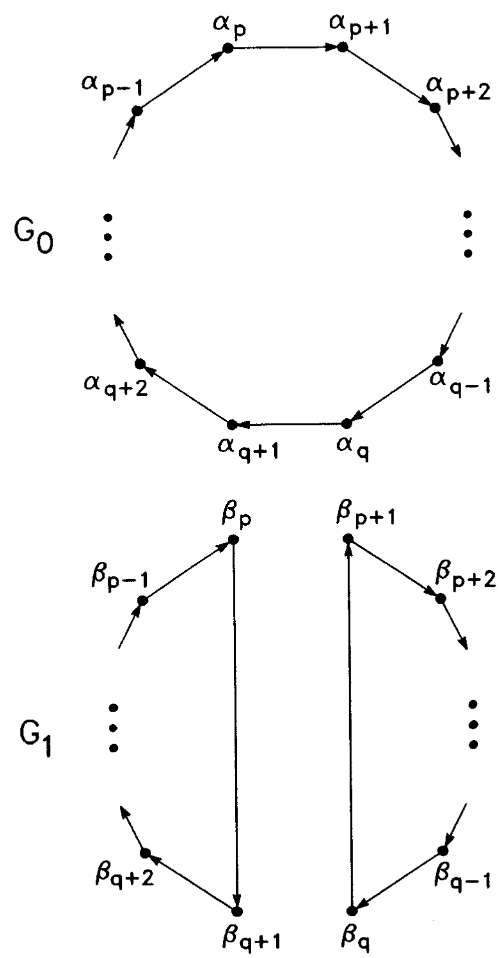


**Figure 1: $G_0$ and $G_1$**

It is instructive to see why the use of the Ajtai-Fagin $(c, r)$-game, as opposed to the original $(c, r)$-game, is important in this proof. The choice of $G_1$ depends on the coloring by the spoiler of $G_0$. Our proof would not work if, as in the original $(c, r)$-game, the duplicator were required to select $G_0$ and $G_1$ before the spoiler colors $G_0$.

Let us consider now the three difficulties that we mentioned in Section 4 about the author's original proof, and see how they have been ameliorated by Fagin, Stockmeyer and Vardi's proof.

**D1:** *The selection of the graphs.* In the original proof, the sizes of the cycles had to be carefully selected. In the new proof, we simply pick $G_0$ to be a sufficiently large cycle.

**D2:** *The duplicator's coloring strategy.* In the new proof, this could not be simpler: the duplicator simply mimics the spoiler's coloring.

**D3:** *The duplicator's pebbling strategy.* In the new proof, we simply appeal to Hanf's condition.

**Example 2: Directed reachability.** Our second example deals with the proof that directed reachability is not in monadic NP. We sketch the idea of the proof, as given by Arora and Fagin [AF94], which is a simplification of the proof of Ajtai and Fagin [AF90].

**Sketch of proof:** Let $\mathcal{S}$ be the class of $(s, t)$-connected graphs. Let $v_1, \ldots, v_n$ be $n$ points, which are used as the set of vertices of the graph $G_0$. The vertex $v_1$ is labeled $s$, and the vertex $v_n$ is labeled $t$. Then $G_0$, the member of $\mathcal{S}$ selected by the duplicator in the Ajtai-Fagin $(c, r)$-game over $\mathcal{S}$, has "forward edges" $\langle v_i, v_{i+1}\rangle$ for $1 \leq i < n$; these form a path from $s$ to $t$. In addition, $G_0$ has certain "backedges" $\langle v_i, v_j\rangle$ where $j < i$. The choice of these backedges are made by probabilistic means; for the proof to work, there cannot be too few or too many backedges. Small cycles are eliminated in the construction by deleting one backedge from each small cycle. We refer to each such graph $G_0$ we obtain as an *$(s, t)$-path with backedges.* If $e$ is one of the forward edges of $G_0$, then denote by $G_0 - e$ the graph that results by deleting the edge $e$. It is clear that (a) there is a path from $s$ to $t$ in $G_0$, but (b) for each forward edge $e$, there is no path from $s$ to $t$ in $G_0 - e$. Thus, $G_0 \in \mathcal{S}$, but $G_0 - e \in \overline{\mathcal{S}}$ for each forward edge $e$. It is now shown probabilistically that for a certain choice of $G_0$ and for each coloring of $G_0$ by the spoiler, a forward edge $e$ can be selected (also probabilistically) so that if (a) $G_1$ is taken to be $G_0 - e$, and (b) the duplicator mimics the coloring of $G_0$ on $G_1$, then Arora and Fagin's condition (Theorem 3.7) is satisfied. Therefore, the duplicator can win the remaining $r$-game. By Theorem 4.3, this is sufficient to show that directed reachability is not in monadic NP. This concludes the proof. ∎

Again, it is instructive to see why the use of the Ajtai-Fagin $(c, r)$-game, as opposed to the original $(c, r)$-game, is important in this proof. If the spoiler knew which edge $e$ were deleted from $G_0$ to form $G_1 = G - e$, this might dramatically influence his coloring of $G_0$ (for example, the spoiler might color the endpoints of $e$ with a special color, say red, not used anywhere else). Then the duplicator would not be able to use the strategy of simply mimicking the spoiler's coloring, since there would be an edge between the red nodes in $G_0$ but not between the red nodes in $G_1$.

Again, let us consider the three difficulties mentioned earlier, and see how this proof helps bypass them.

**D1:** *The selection of the graphs.* In this proof, the graphs are selected by a random procedure, rather than being constructed explicitly. It is shown that with high probability, the duplicator succeeds. Since in particular the probability is nonzero, there exists a winning strategy for the duplicator. This probabilistic approach is potentially very powerful.

**D2:** *The duplicator's coloring strategy.* Once again, the duplicator simply mimics the spoiler's coloring.

**D3:** *The duplicator's pebbling strategy.* Here we simply appeal to Arora and Fagin's condition.

# 6 Some additional comments

Ajtai and Fagin invented the Ajtai-Fagin game because they did not see how to prove that directed reachability is not in monadic NP by using the original game. They posed the question as to whether the same types of graphs they used (a graph $G_0$ that is $(s, t)$-path with backedges, and a graph $G_1$ that is the result of deleting a forward edge $e$ from $G_0$) could have, in principle, been used in the original game to prove the same result. A theorem was proven in [Fag95] that implies that this is indeed the case: in general, the same types of graphs can be used in the original game as in the Ajtai-Fagin game. On the other hand, a more complicated coloring strategy may be required in the original game than the Ajtai-Fagin game. For example, let us consider $G_0$ and $G_1 = G_0 - e$ as above. As we noted earlier, in the original game over these graphs the duplicator cannot simply mimic the spoiler's coloring (as the duplicator did in the Ajtai-Fagin game). This is because, as we noted, the spoiler can color the endpoints of the edge $e$ with a special color.

Finally, we comment on the fact that in the two proofs we discussed in Section 5, the duplicator was able to simply mimic the spoiler's coloring. That is, the duplicator has a winning strategy not just in the Ajtai-Fagin game, but in a game where $G_1$ is required to be colored just like $G_0$. Not surprisingly, the fact that the duplicator has a winning strategy even when $G_1$ must be colored just like $G_0$ corresponds to inexpressibility in a richer logic than that of monadic NP. We now discuss what this richer logic is.

Let $A_1, \ldots, A_k$ be unary relation symbols. Let us define a *color* to be a (formal) conjunction of the form $A'_1 \wedge \ldots\ldots A'_k$, where each $A'_i$ is either $A_i$ or $\neg A_i$. Thus, there are $2^k$ colors. Let $R$ be a (possibly infinite) relation over the natural numbers, with $2^k$ columns. For each such relation $R$ we define a new quantifier $\exists \mathbf{A}_R$, that is interpreted as follows: $\exists \mathbf{A}_R \varphi(A_1, \ldots, A_k)$ holds iff there are unary relations $U_1, \ldots, U_k$ such that $\varphi(A_1, \ldots, A_k)$ holds when $A_1, \ldots, A_k$ are interpreted by $U_1, \ldots, U_k$, respectively, and such that the $2^k$-tuple of the cardinalities of the colors corresponding to $U_1, \ldots, U_k$ is in the relation $R$. Thus, intuitively, $R$ tells the set of possibilities for the distribution of colors. For example, if $R$ consists of the tuples $\langle 2001, 47, 9, 12\rangle, \langle 0, 10000, 96, 4\rangle, \ldots$, and if we call the four colors red, blue, green, and yellow, then, intuitively, $\exists \mathbf{A}_R \varphi$ says "There exists a coloring such that either there are exactly 2001 red points, 47 blue points, 9 green points, and 12 yellow points, or there are exactly 0 red points, 10000 blue points, 96 green points, and 4 yellow points, or there are exactly ... and $\varphi$ holds."

It is not hard to show that the fact that the duplicator has a winning strategy when $G_1$ must be colored just like $G_0$ corresponds to inexpressibility even by sentences of the form $\exists \mathbf{A}_R \psi$, where $\psi$ is first-order.

# 7 Conclusions

As Fagin, Stockmeyer and Vardi observe [FSV95], if we are to make serious progress on resolving difficult problems in computational complexity through using descriptive complexity techniques, we need to develop our descriptive complexity toolkit. Ehrenfeucht-Fraïssé games are the current major tools in our arsenal. In this paper, we discuss some steps that have been taken to ease our task in proving descriptive complexity lower bounds.

In looking back at progress in descriptive complexity, one moral that can be drawn is that we should try to use general principles, rather than ad hoc arguments, in proving inexpressibility results. This makes it easier for others to use these techniques in the future. We also suggest that it would be a useful exercise to go back and look at previous inexpressibility results, to see if general principles can be extracted.

# Acknowledgements

The author is grateful to Lauri Hella, Neil Immerman, Phokion Kolaitis, Clemens Lautemann, Thomas Schwentick, Larry Stockmeyer, Moshe Vardi, and Victor Vianu for helpful comments and suggestions.

# Bibliography


[AF90] M. Ajtai and R. Fagin, Reachability is harder for directed than for undirected finite graphs, *Journal of Symbolic Logic*, **55** (1) March 1990, 113–150.

[AF94] S. Arora and R. Fagin, On winning strategies in Ehrenfeucht-Fraïssé games, Research Report RJ 9833, IBM, June 1994, to appear, *Theoretical Computer Science.*

[Ber76] C. Berge, *Graphs and Hypergraphs*, North-Holland, New York, 1976.

[CFI92] J. Cai and M. Fürer and N. Immerman, An optimal lower bound on the number of variables for graph identification, Combinatorica, **12** 1992, 389–410.

[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.

[Ete95] K. Etessami, Counting Quantifiers, Successor Relations, and Logarithmic Space, *Proc. 10th IEEE Conf. on Structure in Complexity Theory*, 2–11, 1995.

[Fag74] R. Fagin, Generalized first–order spectra and polynomial–time recognizable sets, R. M. Karp, editor, *Complexity of Computation, SIAM-AMS Proceedings*, **7** 1974,43–73.

[Fag75] R. Fagin, Monadic generalized spectra, *Zeitschrift für Mathematische Logik und Grundlagen der Mathematik*, **21** 1975, 89–96.

[Fag83] R. Fagin, Degrees of acyclicity for hypergraphs and relational database schemes, *Journal of the ACM*, **30** (3) July 1983, 514–550.

[Fag93] R. Fagin, Finite-model theory—a personal perspective, *Theoretical Computer Science*, **116** (1) 1993, 3–31.

[Fag95] R. Fagin, Comparing the power of monadic NP games, D. Leivant, editor, *Logic and Computational Complexity*, Lecture Notes in Computer Science, **960**, 414–425, Springer-Verlag, Berlin/NewYork, 1995.

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

[FSV95] R. Fagin and L. Stockmeyer and M. Y. Vardi, On Monadic NP vs. Monadic co-NP, *Information and Computation*, **120** (1) July 1995, 78–92.

[Gai82] H. Gaifman, On local and nonlocal properties, J. Stern, editor, *Logic Colloquium '81*, North Holland, 1982, 105–135.

[Gur84] Y. Gurevich, Toward logic tailored for computational complexity, M. Richter et al., editors, *Computation and Proof Theory*, Lecture Notes in Mathematics, 1104, 175–216, Springer-Verlag, Berlin/New York, 1984.

[Gur90] Y. Gurevich, On finite model theory, S. R. Buss and P. J. Scott, editors, *Perspectives in Computer Science*, Birkhauser,1990. Proceedings of Feasible Mathematics Workshop, Cornell University, June 1989.

[Han65] W. Hanf, Model-Theoretic Methods in the Study of Elementary Logic, J. Addison, L. Henkin and A. Tarski, editors, *The Theory of Models*, North Holland, 1965, 132–145.

[Hel92] L. Hella, Logical hierarchies in PTIME, in *Proc. 7th IEEE Symp. on Logic in Computer Science*, 360–368, 1992.

[IL90] N. Immerman and E. S. Lander, Describing graphs: a first-order approach to graph canonization, A. Selman, editor, *Complexity Theory Retrospective*, 59–81, Springer-Verlag, 1990.

[Imm89] N. Immerman, Descriptive and Computational Complexity, J. Hartmanis, editor, *Computational Complexity Theory, Proc. Symp. Applied Math., Vol. 38*, 75–91, American Mathematical Society, 1989.

[Nur96] M. Nurmonen, On winning strategies with unary quantifiers, *Journal of Logic and Computation*, 1996, to appear.

[Sch94] T. Schwentick, Graph connectivity and monadic NP, *Proc. 35th IEEE Symp. on foundations of Computer Science*, 1994, 614–622.

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

[Sch96] T. Schwentick, On bijections vs. unary functions, *Proc. 13th Symposium on Theoretical Aspects of Computer Science-STACS 96*, to appear.

[Tho91] W. Thomas, On logics, tilings, and automata, *Proc. 18th International Colloq. on Automata, Languages, and Programming*, 441–454, Springer-Verlag, 1991, Lecture Notes in Computer Science 510.