On Reachability in Graphs with Bounded Independence Number

Arfst Nickelsen and Till Tantau*

Technische Universität Berlin
Fakultät für Elektrotechnik und Informatik
10623 Berlin, Germany
{nicke,tantau}@cs.tu-berlin.de

Abstract. We study the reachability problem for finite directed graphs whose independence number is bounded by some constant k. This problem is a generalisation of the reachability problem for tournaments. We show that the problem is first-order definable for all k. In contrast, the reachability problems for many other types of finite graphs, including dags and trees, are not first-order definable. Also in contrast, first-order definability does not carry over to the infinite version of the problem. We prove that the number of strongly connected components in a graph with bounded independence number can be computed using TC^0-circuits, but cannot be computed using AC^0-circuits. We also study the succinct version of the problem and show that it is Π_2^{P}-complete for all k.

1 Introduction

One of the most fundamental problems in graph theory is the reachability problem. For this problem we are asked to decide whether there exists a path from a given source vertex s to a given target vertex t in some graph G. For finite directed graphs this problem, which will be denoted REACH in the following, is well-known to be NL-complete [12, 13]. It is thus easy from a computational point of view and efficient parallel algorithms are known for it. The complexity of the reachability problem drops if we restrict the type of graphs for which we try to solve it. The reachability problem $\mathrm{REACH}_{\mathrm{u}}$ for finite undirected graphs is SL-complete [15] and thus presumably easier to solve. The even more restricted problem $\mathrm{REACH}_{\mathrm{forest}}$ for undirected forests and the problem $\mathrm{REACH}_{\mathrm{out}\le 1}$ for directed graphs in which all vertices have out-degree at most 1 are L-complete [2].

In this paper we study the reachability problem for finite directed graphs whose independence number is bounded by some constant k. The independence number $\alpha(G)$ of a graph G is the maximum number of vertices that can be picked from G such that there is no edge between any two of these vertices. Thus we study the languages $\mathrm{REACH}_{\alpha\le k} := \mathrm{REACH} \cap \big\{\langle G,s,t\rangle \mid \alpha(G) \le k\big\}$ for constant k, where $\langle\,\rangle$ denotes a standard binary encoding. We show that, somewhat surprisingly, $\mathrm{REACH}_{\alpha\le k}$ is first-order definable for all k.

* Work done in part while visiting the University of Rochester, New York. Supported by a TU Berlin Erwin-Stephan-Prize grant.

First-order definability means the following. Let $\tau = \bigl(E^2, s, t\bigr)$ be the signature of directed graphs with two distinguished vertices. The binary relation symbol E represents an edge relation and the constant symbols s and t represent a source and a target vertex. We show that for each k there exists a first-order formula $\phi_{reach,\alpha\leq k}$ over the signature τ for which the following holds: for all *finite* directed graphs $G = (V, E)$ and all $s, t \in V$ the τ-structure (V, E, s, t) is a model of $\phi_{reach,\alpha\leq k}$ iff $\alpha(G) \leq k$ and there is path from s to t in G. The formulas will neither require an ordering on the universe nor the bit predicate [11].

The most prominent examples of graphs with bounded independence number are *tournaments* [18, 20], which are directed graphs with exactly one edge between any two vertices. Their independence number is 1. Conditions for strong connectedness of tournaments (and thus, implicitly, for reachability) were proven in [9], but these conditions yield weaker bounds on the complexity of the reachability problem for tournaments than those shown in the present paper. A different example of graphs with bounded independence number, studied in [4], are directed graphs $G = (V, E)$ whose underlying undirected graph is claw-free, i. e., does not contain the $K_{1,m}$ for some constant m, and whose minimum degree is at least $|V|/3$. Their independence number is at most $3m - 3$.

Languages whose descriptive complexity is first-order are known to be very simple from a computational point of view. They can be decided by a family of AC^0-circuits (constant depth circuits) and also in constant parallel time on concurrent-read, concurrent-write parallel random access machines [16]. Since it is known that L-hard sets cannot be first-order definable [1, 6], $\textsc{reach}_{\alpha\leq k}$ is (unconditionally) easier to solve than $\textsc{reach}$, $\textsc{reach}_{\mathrm{u}}$, and $\textsc{reach}_{\mathrm{forest}}$.

A problem closely related to the reachability problem is the problem of identifying the strongly connected components of a graph. We show that TC^0-circuits (constant depth circuits with threshold gates) can count the strongly connected components in graphs with bounded independence number, but AC^0-circuits cannot—not even in tournaments.

In hardware design one is often concerned with succinctly represented graphs, which are given implicitly via a program or a circuit that decides the edge relation of the graph. Papadimitriou, Yannakakis, and Wagner [19, 23, 24] have shown that the problems $\textsc{succinct-reach}$, $\textsc{succinct-reach}_{\mathrm{u}}$, $\textsc{succinct-reach}_{\mathrm{forest}}$, and $\textsc{succinct-reach}_{\mathrm{out}\leq 1}$ are PSPACE-complete. Opposed to this, we show that $\textsc{succinct-reach}_{\alpha\leq k}$ is Π_2^{P}-complete for all k.

Our results apply only to finite graphs. Let $\textsc{reach}^{\infty}_{\alpha\leq k}$ be the class of all triples (G, s, t) such that G is a (possibly infinite) directed graph with $\alpha(G) \leq k$ in which there is a path from s to t. We show that there does not exist a set of first-order formulas (not even an uncountable one) whose class of models is exactly $\textsc{reach}^{\infty}_{\alpha\leq k}$ for some k.

This paper is organised as follows. In Section 2 we study graph-theoretic definitions and results and prove a general theorem that shows how the independence number of a graph is connected to its different domination numbers. We believe this theorem to be of independent interest. In Section 3 we show that the problem $\textsc{reach}_{\alpha\leq k}$ is first-order definable, by explicitly giving a defining formula. In Section 4 we study the circuit complexity of counting the number of strongly connected components in a graph. In Sections 5 we study the infinite version of our problem and in Section 6 the succinct version.

2 Graph-Theoretic Definitions and Results

In this section we first give definitions of basic graph-theoretic concepts. Then we prove a generalisation of the so-called lion king lemma, see Theorem 2.2. At the end of the section we prove Theorem 2.3, which will be the crucial building block of our first-order definition of $\text{REACH}_{\alpha\leq k}$.

A *graph* is a nonempty set V of vertices together with a set $E \subseteq V \times V$ of directed edges. The *out-degree* of a vertex u is the number of vertices v with $(u, v) \in E$. A *path of length* ℓ in a graph $G = (V, E)$ is a sequence $v_0, \dots, v_\ell$ of vertices with $(v_i, v_{i+1}) \in E$ for $i \in \{0, \dots, \ell - 1\}$. A vertex t is *reachable* from a vertex s if there is a path from s to t. A *strongly connected component* is a maximal vertex set $U \subseteq V$ such that every vertex in U is reachable from every other vertex in U. A set $U \subseteq V$ is an *independent set* if there is no edge in E connecting vertices in U. The maximal size of independent sets in G is its *independence number* $\alpha(G)$. For $i \in \mathbb{N}$, a vertex $u \in V$ is said to *i-dominate* a vertex $v \in V$ if there is a directed path from u to v of length at most i. Let $\text{dom}_i(U)$ denote the set of vertices that are i-dominated by vertices in U. A set $U \subseteq V$ is an *i-dominating set* for G if $\text{dom}_i(U) = V$. The *i-domination number* $\beta_i(G)$ is the minimal size of an i-dominating set for G. A *tournament* is a graph with exactly one edge between any two different vertices and $(v, v) \notin E$ for all $v \in V$. Note that tournaments have independence number 1.

Lemma 2.1. *Let $G = (V, E)$ be a finite graph, $n := |V|$, $\alpha := \alpha(G)$. Then G has at least $\binom{n}{2} \,/\, \binom{\alpha+1}{2}$ edges and there exists a vertex with out-degree at least $(n-1) \,/\, 2\binom{\alpha+1}{2}$.*

Proof. The number of $(\alpha + 1)$-element subsets of V is $\binom{n}{\alpha+1}$. Every such set contains two vertices linked by an edge. Every such edge is in $\binom{n-2}{\alpha-1}$ different $(\alpha+1)$-element subsets of V. Therefore there are at least $\binom{n}{\alpha+1} \,/\, \binom{n-2}{\alpha-1} = \binom{n}{2} \,/\, \binom{\alpha+1}{2}$ edges in G. This also shows that the average out-degree in G is at least $\binom{n}{2} \,/\, n\binom{\alpha+1}{2} = (n-1) \,/\, 2\binom{\alpha+1}{2}$ and one vertex has at least this out-degree. □

Turán [21], referenced in [22], gives an exact formula for the minimal number of edges in a graph as a function of the graph's independence number. However, the simple bound from the above lemma will be more appropriate for our purposes.

Theorem 2.2. *Let $G = (V, E)$ be a finite graph, $n := |V|$, $\alpha := \alpha(G)$. Then $\beta_1(G) \leq \lceil \log_c n \rceil$ and $\beta_2(G) \leq \alpha$, where $c = (\alpha^2 + \alpha)/(\alpha^2 + \alpha - 1)$.*

Proof. We iteratively construct a 1-dominating set D_1 for G of size at most $\lceil \log_c n \rceil$. In each step we put a vertex v_i into D_1 that dominates as many vertices as possible of the subset $V_i \subseteq V$ not dominated so far. Formally, set $V_0 := V$

and for $i \geq 1$, as long as V_{i-1} is not empty, choose a vertex $v_i \in V_{i-1}$ such that $V_i := V_{i-1} \setminus \mathrm{dom}_1(\{v_i\})$ is as small as possible. Let $i_{\max}$ be the first i such that V_i is empty. By Lemma 2.1 the out-degree of v_i is at least $(|V_{i-1}| - 1) \,/\, 2\binom{\alpha+1}{2}$ and thus

$$\begin{aligned} |V_i| &\leq |V_{i-1}| - 1 - \frac{|V_{i-1}| - 1}{2\binom{\alpha+1}{2}} < |V_{i-1}| - \frac{|V_{i-1}|}{2\binom{\alpha+1}{2}} \\ &= |V_{i-1}|\left(1 - \frac{1}{2\binom{\alpha+1}{2}}\right) = |V_{i-1}|\left(\frac{\alpha^2+\alpha-1}{\alpha^2+\alpha}\right) = \frac{|V_{i-1}|}{c}. \end{aligned}$$

This shows that the size of V_i decreases by at least the factor c in each step. Thus after at most $\lceil \log_c n \rceil$ iterations the set V_i is empty and $D_1 := \{v_1, \dots, v_{i_{\max}}\}$ is the desired 1-dominating set.

We next construct a 2-dominating set D_2 of size at most α by removing superfluous vertices from D_1. Formally, let $W_{i_{\max}} := \{v_{i_{\max}}\}$ and let $W_{i-1} := W_i$ if $v_i \in \mathrm{dom}_1(W_i)$, and $W_{i-1} := W_i \cup \{v_i\}$ otherwise. Clearly, $D_2 := W_1$ is a 2-dominating set. To prove $|D_2| \leq \alpha$, assume that D_2 contains at least $\alpha + 1$ vertices $v_{i_1}, \dots, v_{i_{\alpha+1}} \in D_1$. Since these vertices cannot be independent, there must exist indices i_r and i_s such that $(v_{i_r}, v_{i_s}) \in E$. By construction of the set D_1, this can only be the case if $i_s > i_r$. But then $v_{i_r} \notin D_2$ by construction of W_{i_r}, a contradiction. □

For tournaments G, Theorem 2.2 yields $\beta_1(G) \leq \log_2(n)$ and $\beta_2(G) = 1$. The first result was first proved by Megiddo and Vishkin in [17], where it was used to show that the dominating set problem for tournaments is not NP-complete, unless $\mathrm{NP} \subseteq \mathrm{DTIME}(n^{O(\log n)})$. The second result is also known as the lion king lemma, which was first noticed by Landau in [14] in the study of animal societies, where the dominance relations on prides of lions form tournaments. It has applications in the study of P-selective sets [10] and many other fields.

Theorem 2.3. *Let $G = (V, E)$ be a finite graph, $n := |V|$, $\alpha := \alpha(G)$, $c := (\alpha^2+\alpha)/(\alpha^2+\alpha-1)$, and $s, t \in V$. Then the following statements are equivalent:*

1. *There is no path from s to t in G.*
2. *There is a subset $D_1 \subseteq V$ with $|D_1| \leq \lceil \log_c n \rceil$ such that $\mathrm{dom}_1(D_1)$ is closed under reachability, $s \in \mathrm{dom}_1(D_1)$ and $t \notin \mathrm{dom}_1(D_1)$.*
3. *There is a subset $D_2 \subseteq V$ with $|D_2| \leq \alpha$ such that $\mathrm{dom}_2(D_2)$ is closed under reachability, $s \in \mathrm{dom}_2(D_2)$ and $t \notin \mathrm{dom}_2(D_2)$.*

Proof. Both 2 and 3 imply 1, since no path starting at a vertex s inside a set that closed is under reachability can 'leave' this set to arrive at a vertex t outside this set. To show that 1 implies 2, consider the set S of vertices reachable from s in G. Then S is closed under reachability, $s \in S$ and $t \notin S$. The induced graph $G' := (S, E \cap (S \times S))$ also has independence number at most α. Therefore, by Theorem 2.2, the graph G' has a 1-dominating set D_1 of size at most $\lceil \log_c n \rceil$. To show that 1 implies 3, consider the same graph G' once more. By Theorem 2.2 it also has a 2-dominating set D_2 of size at most α. □

3 First-Order Definability of the Problem

In this section we show that reachability in graphs with bounded independence number is first-order definable. We start with a review of some basic notions from descriptive complexity theory.

We use the *signature* or *vocabulary* $\tau = \big(E^2, s, t\big)$. It consists of a binary relation symbol E, representing an edge relation, and constant symbols s and t, representing a source and a target vertex. A *τ-structure* is a tuple (V, E, s, t) such that $E \subseteq V \times V$ and $s, t \in V$. We do not distinguish notationally between the symbols in the signature and their interpretation in a structure, because it is always clear from the context which of the two meanings is intended. The standardised binary code of a finite τ-structure (V, E, s, t) will be denoted $\langle V, E, s, t\rangle$. A set A of codes of finite τ-structures is *first-order definable* if there exists a first-order formula ϕ over the signature τ such that for all finite τ-structures (V, E, s, t) we have $(V, E, s, t) \models \phi$ iff $\langle V, E, s, t\rangle \in A$.

Theorem 3.1. *For each k, $\text{REACH}_{\alpha \le k}$ is first-order definable.*

Proof. Let $k \ge 1$ be fixed. We give a stepwise construction of a formula $\phi_{reach,\alpha\le k}$ such that $(V, E, s, t) \models \phi_{reach,\alpha\le k}$ iff $\langle V, E, s, t\rangle \in \text{REACH}_{\alpha\le k}$. Roughly spoken, the formula $\phi_{reach,\alpha\le k}$ will say '$\alpha(G) \le k$ and it is not the case that condition 3 of Theorem 2.3 holds for s and t'.

Let $\phi_{distinct}(v_1, \dots, v_k) \equiv \bigwedge_{i\ne j}[v_i \ne v_j]$. This formula expresses that vertices are distinct. The property '$\alpha(G) \le k$' can be expressed as follows:

$$\phi_{\alpha\le k} \equiv (\forall v_1, \dots, v_{k+1})\Big[\phi_{distinct}(v_1, \dots, v_{k+1}) \to \bigvee\nolimits_{i\ne j} E(v_i, v_j)\Big].$$

The next two formulas express that a vertex v, respectively a set $\{v_1, \dots, v_m\}$ of vertices, 2-dominates a vertex u:

$$\phi_{\textit{2-dom}}(v, u) \equiv v = u \vee E(v, u) \vee (\exists z)\big[E(v, z) \wedge E(z, u)\big],$$
$$\phi_{\textit{2-dom}}(v_1, \dots, v_m, u) \equiv \phi_{\textit{2-dom}}(v_1, u) \vee \dots \vee \phi_{\textit{2-dom}}(v_m, u).$$

Since $\beta_2(G) \le \alpha(G) \le k$, condition 3 of Theorem 2.3 can be expressed as follows:

$$\begin{aligned}&\phi_{condition} \equiv (\exists v_1, \dots, v_k)\\ &\quad\Big[\phi_{\textit{2-dom}}(v_1, \dots, v_k, s) \wedge \neg\phi_{\textit{2-dom}}(v_1, \dots, v_k, t) \wedge\\ &\quad(\forall u, v)\big[\big(\phi_{\textit{2-dom}}(v_1, \dots, v_k, u) \wedge \neg\phi_{\textit{2-dom}}(v_1, \dots, v_k, v)\big) \to \neg E(u, v)\big]\Big].\end{aligned}$$

The desired formula $\phi_{reach,\alpha\le k}$ is given by $\phi_{\alpha\le k} \wedge \neg\phi_{condition}$. □

Note that the formula $\phi_{reach,\alpha\le k}$ constructed in the proof has quantifier alternation depth three, beginning with a universal quantifier.

Theorem 3.1 be easily extended to the following larger class of graphs: define the *r-independence number* $\alpha_r(G)$ of a graph G as the maximal size of an r-independent set in G, which is a vertex subset such that there is no path of length at most r between any two different vertices in this subset. Then reachability in graphs with $\alpha_r(G) \le k$ is first-order definable for all $k, r \in \mathbb{N}$.

4 Circuit Complexity of the Problem

In this section we study the circuit complexity of the problem $\text{REACH}_{\alpha\leq k}$, as well as the complexity of counting the number of strongly connected components in a graph with bounded independence number. We show that this number can be computed using TC^0-circuits, but cannot be computed using AC^0-circuits.

A family $\mathcal{C} = (C_n)_{n\in\mathbb{N}}$ of circuits is a *family of* AC^0*-circuits* if each C_n has n input gates, their size is bounded by a polynomial in n, their depth is bounded by a constant, and each C_n consist of unbounded fan-in/fan-out and-, or-, and not-gates. For TC^0-circuits we also allow threshold gates, whose output is 1 if the number of 1's at the input exceeds some threshold. For $x \in \{0,1\}^n$ we write $\mathcal{C}(x)$ for the output produced by C_n on input x. The output may be a bitstring since we allow multiple output gates. A circuit family $\mathcal{C}$ *decides* a set $A \subseteq \{0,1\}^*$, respectively *computes* a function $f\colon \{0,1\}^* \to \{0,1\}^*$, if for all $x \in \{0,1\}^*$ we have $x \in A$ iff $\mathcal{C}(x) = 1$, respectively $f(x) = \mathcal{C}(x)$.

As shown by Lindell [16], every first-order definable set can be decided by AC^0-circuits. In particular, by Theorem 3.1 there exists, for each k, an AC^0-circuit family $\mathcal{C}^k$ that decides $\text{REACH}_{\alpha\leq k}$. We now sketch how these families can be used to decrease the average case complexity of REACH, which is L-hard and thus does not have AC^0-circuits [1, 6]. Suppose there exists a constant k for which we expect $\alpha(G) \leq k$ to hold with high probability for input graphs G. Then whenever $\alpha(G) \leq k$ holds, we can use $\mathcal{C}^k$ to decide in constant depth whether there is a path from s to t. For graphs with $\alpha(G) > k$ we use a slow standard reachability circuit to decide whether such a path exists. If the probability of $\alpha(G) \leq k$ is sufficiently large, the preprocessing will decrease the average time taken by the circuit to produce its output.

A problem closely related to the reachability problem is the problem of counting strongly connected components. The following theorem pinpoints the exact circuit complexity of this counting problem for graphs with bounded independence number. Let $\zeta_{\alpha\leq k}\colon \{0,1\}^* \to \{0,1\}^*$ be the function that maps the code $\langle G\rangle$ of a graph G to the binary representation of the number of strongly connected components in G if $\alpha(G) \leq k$, and that maps $\langle G\rangle$ to 0 if $\alpha(G) > k$.

Theorem 4.1. *For each k, $\zeta_{\alpha\leq k}$ can be computed by* TC^0*-circuits, but not by* AC^0*-circuits.*

Proof. Let k be fixed. Let $\phi_{reach,\alpha\leq k}(u,v)$ be the formula with two free variables expressing that v is reachable from u and that the underlying graph has independence number at most k. It is obtained from $\phi_{reach,\alpha\leq k}$ from the proof of Theorem 3.1 by replacing the constant symbols s and t by variables u and v. Consider the formula

$$\phi_{rep}(v) \equiv (\forall u)\big[u < v \to \big(\neg\phi_{reach,\alpha\leq k}(u,v) \vee \neg\phi_{reach,\alpha\leq k}(v,u)\big)\big],$$

where '$<$' is a relation that is interpreted as a total ordering of the set of vertices.

For a graph G with $\alpha(G) \leq k$, the formula $\phi_{rep}(v)$ will be true exactly for the smallest members (with respect to the ordering $<$) of each strongly connected component. Thus, the number of vertices v for which $\phi_{rep}(v)$ holds is exactly the number of strongly connected components in G. Since ϕ_{rep} is a first-order formula, there exists a family of AC^0-circuits that maps $\langle\{v_1,\ldots,v_n\},E\rangle$ to a bitstring in which the i-th position is 1 iff $(\{v_1,\ldots,v_n\},E) \models \phi_{rep}(v_i)$. Since the number of 1's in this bitstring can be computed in constant depth using threshold gates, $\zeta_{\alpha\leq k}$ can be computed by TC^0-circuits.

Next, for the sake of contradiction, assume that there exists an AC^0-circuit family $\mathcal{C}$ that computes $\zeta_{\alpha\leq k}$. We construct an AC^0-circuit for the parity function, contradicting the results of Ajtai et al. [1, 6]. Let a bitstring $b = b_1 \ldots b_n$ be given as input. Define a tournament $G = (\{1,\ldots,n+1\},E)$ as follows: for $i+1 < j$ there is an edge from j to i; for $i+1 = j$ there is an edge from j to i if $b_i = 1$; otherwise there is an edge from i to j. If b contains no 1's, the tournament will form one big circle, thus having just one strongly connected component. Every additional 1 in b adds one strongly connected component. The parity of b is thus given by the toggled least-significant bit of $\mathcal{C}(\langle G\rangle)$. □

5 Infinite Version of the Problem

In this section we study the class $\textsc{reach}^{\infty}_{\alpha\leq k}$ and show that the results of Section 3 on the first-order definability of $\textsc{reach}_{\alpha\leq k}$ do not carry over to $\textsc{reach}^{\infty}_{\alpha\leq k}$. This class contains all triples (G,s,t) such that G is a (possibly infinite) graph with $\alpha(G) \leq k$ in which there is a path from s to t. We start with a review of the relevant notions from model theory.

Let τ be a signature. A class K of τ-structures is called *elementary (over finite structures)* if there exists a first-order formula ϕ over τ such that for every (finite) τ-structure $\mathcal{A}$ we have $\mathcal{A} \models \phi$ iff $\mathcal{A} \in K$. (Some authors use 'finitely axiomatisable' instead of 'elementary'.) A class K of τ-structures is *Δ-elementary* if there exists a set Φ of first-order formulas over τ such that for every τ-structure $\mathcal{A}$ we have $\mathcal{A} \models \Phi$ iff $\mathcal{A} \in K$.

Fact 5.1 (Compactness Theorem). *Let Φ be a set of first-order formulas such that every finite $\Phi_0 \subseteq \Phi$ has a model. Then Φ has a model.*

With these definitions, Theorem 3.1 simply states that $\textsc{reach}^{\infty}_{\alpha\leq k}$ is elementary over finite structures for all k. The below proof that $\textsc{reach}^{\infty}_{\alpha\leq k}$ is not even Δ-elementary follows the standard pattern of proofs applying the compactness theorem. The only essential part is the construction of appropriate model graphs for finite subsets of a hypothetical axiomatisation of $\textsc{reach}^{\infty}_{\alpha\leq k}$.

Theorem 5.2. $\textsc{reach}^{\infty}_{\alpha\leq k}$ *is not Δ-elementary for any k.*

Proof. Assume that there exists a set Φ of first-order formulas with $(V,E,s,t) \models \Phi$ iff $(V,E,s,t) \in \textsc{reach}^{\infty}_{\alpha\leq k}$. For each $n \in \mathbb{N}$ define the following formula ψ_n, which is fulfilled by a graph iff there is a path of length n from s to t.

$$\psi_n \equiv (\exists v_1,\ldots,v_{n-1})\big[E(s,v_1) \wedge E(v_1,v_2) \wedge \cdots \wedge E(v_{n-2},v_{n-1}) \wedge E(v_{n-1},t)\big].$$

Consider the set $\Psi := \Phi \cup \{\neg\psi_1, \neg\psi_2, \neg\psi_3, \ldots\}$. We claim that every finite $\Psi_0 \subseteq \Psi$ has a model (V, E, s, t). To see this, let n be large enough such that for all $i \geq n$ we have $\neg\psi_i \notin \Psi_0$ and define a graph $G = (V, E)$ by $V := \{1, \ldots, n+1\}$ and $(i, j) \in E$ iff $j \leq i + 1$. Then $\alpha(G) = 1 \leq k$ and the shortest path from $s := 1$ to $t := n + 1$ has length n. Thus (V, E, s, t) is a model of Ψ_0.

Since every finite subset of Ψ has a model, Ψ has a model (V, E, s, t) by the compactness theorem. Since this model fulfills $\neg\psi_n$ for all n, there can be no path of finite length from s to t in $G = (V, E)$. Thus Φ has a model that is not an element of $\textsc{reach}^{\infty}_{\alpha \leq k}$. □

6 Succinct Version of the Problem

In this section we study succinctly represented graphs. Such graphs are given implicitly via a description in some description language. Since succinct representations allow one to encode large graphs into small codes, checking properties is (provably) harder for succinctly represented graphs than for graphs coded in the usual way. Papadimitriou et al. [19, 24] have shown that most interesting problems for succinctly represented graphs are PSPACE-complete or even NEXP-complete. The following formalisation of succinct graph representations is due to Galperin and Wigderson [7], but others are also possible [24, 8].

Definition 6.1. *A* succinct representation *of a graph* $G = \big(\{0,1\}^n, E\big)$ *is a* $2n$*-input circuit* C *such that for all* $u, v \in \{0,1\}^n$ *we have* $(u, v) \in E$ *iff* $C(uv) = 1$.

The circuit tells us for any two vertices of the graph whether there is a directed edge between them or not. Note that there is no need to bound the size of C.

Definition 6.2. *Let* $A \subseteq \{\langle G, s, t\rangle \mid G = (V, E)$ *is a finite graph,* $s, t \in V\}$*. Then* $\textsc{succinct}$-A *is the set of all codes* $\langle C, s, t\rangle$ *such that* C *is a succinct representation of a graph* G *with* $\langle G, s, t\rangle \in A$*.*

Theorem 6.3. *For each* k*,* $\textsc{succinct-reach}_{\alpha \leq k}$ *is* Π^{P}_2*-complete.*

Proof. We first show $\textsc{succinct-reach}_{\alpha \leq k} \in \Pi^{\mathrm{P}}_2$. Let $\langle C, s, t\rangle$ be an input and let C represent a graph $G = (V, E)$ with $V = \{0,1\}^n$. Note that $\log_2 |V| = n$. We first check whether $\alpha(G) \leq k$, which can easily be done using a coNP-machine. We then check whether there is path from s to t in G. By Theorem 2.3 this is case iff *for all* sets $D_1 \subseteq \{0,1\}^n$ of size at most $\beta_1(G)$ either $s \notin \mathrm{dom}_1(D_1)$ or $t \in \mathrm{dom}_1(D_1)$ or $\mathrm{dom}_1(D_1)$ is not closed under reachability, i.e., *there exist* vertices $u \in \mathrm{dom}_1(D_1)$ and $v \in \{0,1\}^n \setminus \mathrm{dom}_1(D_1)$ such that $C(uv) = 1$. Since $\beta_1(G) \leq \lceil \log_c 2^n \rceil \leq \lceil n/\log_2 c \rceil$, the size of the D_1's that need to be checked is linear in n. Thus the 'for all … exists … ' test is a Π^{P}_2-algorithm, since a membership test for the set $\mathrm{dom}_1(D_1)$ can be performed in polynomial time.

We now prove that even the reachability problem $\textsc{succinct-reach}_{\mathrm{tourn}}$ for tournaments is Π^{P}_2-hard. Let $L \in \Pi^{\mathrm{P}}_2$ be any language. By the quantifier characterisation of the polynomial hierarchy [25] there exists a polynomial time decidable ternary relation R and a constant c such that

$$L = \big\{x \mid (\forall y, |y| = |x|^c)(\exists z, |z| = |x|^c)\big[R(x, y, z)\big]\big\}.$$

We construct a reduction from L to SUCCINCT-REACH$_{\text{tourn}}$. On input x we construct, in polynomial time, a circuit C and two bitstrings s, t such that $x \in L$ iff $\langle C, s, t\rangle \in$ SUCCINCT-REACH$_{\text{tourn}}$. Let n denote the length of x and let $\ell := n^c$.

The circuit C will represent a highly structured tournament G of exponential size. The vertex set of G is $V = \{0,1\}^{2\ell+1}$. Each vertex $v \in V$ can be split into a 'y-component' $y \in \{0,1\}^{\ell+1}$ and a 'z-component' $z \in \{0,1\}^{\ell}$ with $yz = v$. All vertices that have the same y-component form a *level*. All vertices on the same level are connected such that they form a strongly connected subtournament of G. We say a level is *above* another level if its y-component is lexicographically larger than the other level's y-component.

Edges between different levels generally point 'downwards', i. e., from higher levels to lower levels. The only exception are edges between a vertex with y-component $0\tilde{y}$ with $\tilde{y} \in \{0,1\}^{\ell}$ and the vertex with the same z-component on the level directly above. Such an edge points 'upwards' iff $R(x, \tilde{y}, z)$. The source is any vertex on the bottom level, the target is any vertex on level 10^{ℓ}.

The graph G is a tournament and the representing circuit C can be constructed in polynomial time. From each level y one can go (at best) only one level higher to the next level y', since all edges between non-neighbouring levels point downwards. Since all vertices on the same level are connected, if one can reach a vertex v on level $0\tilde{y}$, one can reach a vertex on the level directly above iff $R(x, y', z)$ holds for some $z \in \{0,1\}^{\ell}$. So in order to get from the source to the target, for all $\tilde{y} \in \{0,1\}^{\ell}$ there must exist a $z \in \{0,1\}^{\ell}$ such that $R(x, \tilde{y}, z)$. □

7 Conclusion and Open Problems

We showed that the complexity of the reachability problem for graphs with bounded independence number is lower than the complexity of the corresponding problem for, say, forests. However, we did not claim that is also easier to *actually find a path* in a tournament. While it is easily seen that there is a function in FL that maps every forest to a path from the first to the last vertex, provided such a path exists, we do not know whether such a function exists for tournaments. We recommend this problem for further research.

We do not know whether the three levels of quantifier alternation in the first-order formula for REACH$_{\alpha \le k}$ are necessary, but conjecture that this is the case. Since we do not refer to an ordering relation in our first-order formula, it seems promising to use an Ehrenfeucht-Fraïssé game [3, 5] to prove this.

In the succinct setting, we proved that the problem SUCCINCT-REACH$_{\alpha \le k}$ is Π_2^{P}-complete for all k. Opposed to this, for $r > 1$ our arguments only show SUCCINCT-REACH$_{\alpha_r \le k} \in \Pi_3^{\text{P}}$. In particular, we would like to know the exact complexity of SUCCINCT-REACH$_{\alpha_2 \le 1}$.

Acknowledgments. We would like to thank Mitsu Ogihara, Ken Regan, Alan Selman and Leen Torenvliet for helpful discussions.

References

1. M. Ajtai. Σ_1^1 formulae on finite structures. *Annals of Pure and Applied Logic*, 24:1–48, 1983.
2. S. Cook and P. McKenzie. Problems complete for deterministic logarithmic space. *J. Algorithms*, 8(3):385–394, 1987.
3. A. Ehrenfeucht. An application of games to the completeness problem for formalized theories. *Fundamenta Mathematicae*, 49:129–141, 1961.
4. R. Faudree, R. Gould, L. Lesniak, and T. Lindquester. Generalized degree conditions for graphs with bounded independence number. *J. Graph Theory*, 19(3):397–409, 1995.
5. R. Fraïssé. Sur quelques classifications des systèmes de relations. *Publ. Sci. Univ. Alger. Sér. A*, 1:35–182, 1954.
6. M. Furst, J. Saxe, and M. Sipser. Parity, circuits, and the polynomial-time hierarchy. *Math. Systems Theory*, 17(1):13–27, 1984.
7. H. Galperin and A. Wigderson. Succinct representations of graphs. *Inform. Control*, 56(3):183–198, 1983.
8. G. Gottlob, N. Leone, and H. Veith. Succinctness as a source of complexity in logical formalisms. *Annals of Pure and Applied Logic*, 97(1–3):231–260, 1999.
9. F. Harary and L. Moser. The theory of round robin tournaments. *Amer. Math. Monthly*, 73:231–246, 1966.
10. L. Hemaspaandra and L. Torenvliet. Optimal advice. *Theoretical Comput. Sci.*, 154(2):367–377, 1996.
11. N. Immerman. *Descriptive Complexity*. Springer-Verlag, 1998.
12. N. Jones. Space-bounded reducibility among combinatorial problems. *J. Comput. Syst. Sci.*, 11(1):68–85, 1975.
13. N. Jones, Y. Lien, and W. Laaser. New problems complete for nondeterministic log space. *Math. Systems Theory*, 10:1–17, 1976.
14. H. Landau. On dominance relations and the structure of animal societies, III: The condition for secure structure. *Bull. Mathematical Biophysics*, 15(2):143–148, 1953.
15. H. Lewis and C. Papadimitriou. Symmetric space-bounded computation. *Theoretical Comput. Sci.*, 19(2):161–187, 1982.
16. S. Lindell. A purely logical characterization of circuit uniformity. In *Proc. 7th Struc. in Complexity Theory Conf.*, pages 185–192. IEEE Computer Society, 1992.
17. N. Megiddo and U. Vishkin. On finding a minimum dominating set in a tournament. *Theoretical Comput. Sci.*, 61(2–3):307–316, 1988.
18. J. Moon. *Topics on Tournaments*. Holt, Rinehart, and Winston, 1968.
19. C. Papadimitriou and M. Yannakakis. A note on succinct representations of graphs. *Inform. Control*, 71(3):181–185, 1986.
20. K. Reid and L. Beineke. Tournaments. In *Selected Topics in Graph Theory*, pages 169–204. Academic Press, 1978.
21. P. Turán. Eine Extremalaufgabe aus der Graphentheorie (in Hungărian). *Matem. és Physikai Lapok*, 48:436–452, 1941.
22. P. Turán. On the theory of graphs. *Colloquium Math.*, 3:19–30, 1954.
23. K. Wagner. The complexity of problems concerning graphs with regularities. In *Proc. 7th Symposium on Math. Foundations of Comp. Sci.*, volume 176 of *Lecture Notes in Computer Science*, pages 544–552. Springer-Verlag, 1984.
24. K. Wagner. The complexity of combinatorial problems with succinct input representation. *Acta Informatica*, 23(3):325–356, 1986.
25. C. Wrathall. Complete sets and the polynomial-time hierarchy. *Theoretical Comput. Sci.*, 3(1):23–33, 1976.