A Note on the Complexity of Directed Clique

Grzegorz Gutowski
Institute of Theoretical Computer Science, Faculty of Mathematics and Computer Science, Jagiellonian University, Kraków, Poland

Mikołaj Rams
Institute of Theoretical Computer Science, Faculty of Mathematics and Computer Science, Jagiellonian University, Kraków, Poland

Abstract

For a directed graph G, and a linear order $\ll$ on the vertices of G, we define the *backedge graph* $G^{\ll}$ to be the undirected graph on the same vertex set with edge $\{u, w\}$ in $G^{\ll}$ if and only if (u, w) is an arc in G and $w \ll u$. The *directed clique number* of a directed graph G is defined as the minimum size of the maximum clique in the backedge graph $G^{\ll}$ taken over all linear orders $\ll$ on the vertices of G. A natural computational problem is to decide for a given directed graph G and a positive integer t, if the directed clique number of G is at most t. This problem has polynomial algorithm for $t = 1$ and is known to be NP-complete for every fixed $t \geq 3$, even for tournaments. In this note we prove that this problem is Σ_2^{P}-complete when t is given on the input.

2012 ACM Subject Classification Theory of computation

Keywords and phrases Directed Clique, Computational Complexity, Polynomial Hierarchy

Digital Object Identifier 10.4230/LIPIcs.WG.2026.22

Funding *Grzegorz Gutowski*: Partially supported by grant no. 2023/49/B/ST6/01738 from National Science Centre, Poland.

1 Introduction

All graphs discussed in this paper are finite, simple, and loopless. For directed graphs we allow for parallel arcs in opposite directions. We do not allow for parallel arcs in the same direction. The vertex set and the edge set of an undirected graph G are denoted by $V(G)$ and $E(G)$. The vertex set and the arc set of a directed graph G are denoted by $V(G)$ and $A(G)$. For a subset $U \subseteq V(G)$, $G[U]$ denotes the subgraph of G *induced* by U. A directed graph T is a *tournament* when for each pair $u, w \in V(G)$ of vertices, exactly one of the (u, w) or (w, u) is an arc in $A(G)$. For a directed graph G, and a linear order $\ll$ on the vertex set $V(G)$ we define the set of *backedges* to be:

$$\mathrm{BE}(G, \ll) = \{\{u, w\} : (u, w) \in A(G) \wedge w \ll u\}.$$

We define the *backedge graph* $G^{\ll}$ to be the undirected graph with vertex set $V(G)$ and edge set $\mathrm{BE}(G, \ll)$. We say that a directed graph G is *acyclic* if it does not contain a directed cycle. The *transitive tournament* TT_n with n vertices is the only tournament with n vertices that is acyclic.

The *complete directed graph* Q_n with n vertices is the directed graph with n vertices and all possible $n(n-1)$ arcs. An undirected graph with t vertices and all possible $\binom{t}{2}$ edges is a *clique* K_t. We say that an undirected graph G *contains* a clique K_t, if there is a subset $U \subseteq V(G)$ of t vertices such that $G[U]$ is a clique K_t. For a G that does not contain a clique K_t, we say that G is *K_t-free*. The *clique number* of an undirected graph G, denoted $\omega(G)$, is the maximum t such that G contains a clique K_t. The natural computational problem is

52nd International Workshop on Graph-Theoretic Concepts in Computer Science (WG 2026).
Editors: Jan Goedgebeur and Paweł Rzążewski; Article No. 22; pp. 22:1–22:10
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

Problem:	CLIQUE
Input:	An undirected graph G and a positive integer t
Question:	Yes if and only if $\omega(G) \geq t$

This problem has broad practical applications in different areas of discrete optimization and it is also of great theoretical interest. It is one of the first problems known to be NP-complete, see [5], and is used as a base for countless reductions.

Following the work of Kim [6], Nguyen, Seymour and Scott [8], Aboulker, Aubian, Charbit and Lopes [1] we define the *directed clique number* of a directed graph G, denoted $\vec{\omega}(G)$, as the minimum number t such that there exists linear order $\ll$ on $V(G)$ with the clique number of the backedge graph $G^{\ll}$ equal to t, i.e.

$$\vec{\omega}(G) = \min_{\ll} \omega(G^{\ll}),$$

and the natural computational problem is

Problem:	DIRECTEDCLIQUE
Input:	A directed graph G and a positive integer t
Question:	Yes if and only if $\vec{\omega}(G) \leq t$

Nguyen, Seymour and Scott [8, Section 6] were focused on tournaments and asked the following question. *„Can we test whether a tournament has small directed clique number (even in an approximate sense) in polynomial time? Is it in* co-NP*?"* Let us define a variant of the directed clique problem in which the input graph is required to be a tournament.

Problem:	TOURNAMENTCLIQUE
Input:	A tournament T and a positive integer t
Question:	Yes if and only if $\vec{\omega}(T) \leq t$

We also define parametrized variants t-DIRECTEDCLIQUE and t-TOURNAMENTCLIQUE in which t is some fixed value instead of being given on the input.

It is an easy observation that $\vec{\omega}(G) = 0$ iff G is an empty graph, and that $\vec{\omega}(G) = 1$ iff G is acyclic. Acyclicity can be easily checked using depth first search algorithm, and we get that 1-DIRECTEDCLIQUE and 1-TOURNAMENTCLIQUE are in P. For every $t \geq 3$, Aubian [4] showed that both t-TOURNAMENTCLIQUE and t-DIRECTEDCLIQUE are NP-complete and conjectured that the same holds for $t = 2$. To our best knowledge, this question has not been resolved yet.

Other graph parameters can also be defined for directed graphs using the backedge graph similarly as for the definition of the directed clique number. See the discussion of other intersting parameters in a paper on degree-width of directed graphs by Aboulker, Oijid, Petit, Rocton and Simon [2]. The classical notion of *dichromatic number*, defined first by Neumann-Lara [7], can equivalently be defined for a directed graph G as the minimum chromatic number of $G^{\ll}$ taken over all linear orders $\ll$ on $V(G)$. Therefore, out of these similarly defined parameters, the dichromatic number attracts the most attention. Not surprisingly, the dichromatic number and the directed clique number are related, see [8, 1, 4, 2].

Nguyen, Seymour and Scott [8] asked if TOURNAMENTCLIQUE is in P or co-NP. We are unable to answer this questions, but we provide some strong evidence that the answer should be negative. We do that by settling the complexity of the more general problem on directed graphs. Observe that the problem DIRECTEDCLIQUE on instance $\langle G, t\rangle$ is naturally expressed as:

$$\exists_{\ll \text{ - a linear order on } V(G)} : \quad \forall_{A \subseteq V(G), |A| = t+1} : \quad A \text{ does not induce a } (t+1)\text{-clique in } G^{\ll},$$

and from this formulation we easily get that DIRECTEDCLIQUE is in the second level of polynomial hierarchy class Σ_2^{P}. Consult textbook by Arora and Barak [3, Chapter 5] for an introduction of the polynomial hierarchy. Schaefer and Umans [9, 10, 11] give an extensive list of complete problems for different classes in the polynomial hierarchy. For a very brief introduction, Σ_2^{P} is defined as $\mathsf{NP}^{\mathsf{NP}}$ – a class of languages decidable in polynomial time by nondeterministic Turing machines witch access to an NP-oracle, where an NP-oracle allows to test any language in NP in a single step of execution. The canonical complete problem for Σ_2^{P} is the following.

Problem: EXISTENTIAL-2-LEVEL-SAT
Input: Boolean formula $\varphi(x_1, \ldots, x_a, y_1, \ldots, y_b)$ with variables in two disjoint sets $\{x_1, \ldots, x_a\}$ and $\{y_1, \ldots, y_b\}$
Question: Yes if and only if the following Boolean formula is true.
$\exists_{x_1,x_2,\ldots,x_a} : \quad \forall_{y_1,y_2,\ldots,y_b} : \quad \varphi(x_1, \ldots, x_a, y_1, \ldots, y_b)$

It was independently proved by Stockmeyer [12] and Wrathall [13] that the class Σ_2^{P} is exactly the class of languages reducible to EXISTENTIAL-2-LEVEL-SAT via polynomial-time many-one reductions. This fact makes Σ_2^{P} a natural complexity class for DIRECTEDCLIQUE and TOURNAMENTCLIQUE problems. We make the following conjecture that would settle the complexity of both problems.

▶ **Conjecture 1.** *TOURNAMENTCLIQUE is Σ_2^{P}-complete.*

An obstacle in proving Conjecture 1 is the lack of constructions of tournaments with high directed clique number. Aboulker, Aubian, Charbit and Lopes [1] give a simple construction of tournaments with $\vec{\omega}(T) \geq \Omega(\log |V(T)|)$. We are unaware of any construction of tournaments with higher directed clique number. If Conjecture 1 is true, then we should expect that there are tournaments with directed clique number polynomial in the number of vertices. To complement the logarithmic lower bounds, we provide a polynomial upper bound on the directed clique number. In Section 3 we give a proof that for any tournament T we have $\vec{\omega}(T) \leq \sqrt{2\,|V(T)|}$. Let us note that for a random order $\ll$ on the vertices of TT_n we expect the clique number of $G^{\ll}$ to be $\Theta(\sqrt{V(T)})$. Nevertheless, $\vec{\omega}(TT_n) = 1$. We conjecture that the lower bound for the directed clique number of tournaments can be improved significantly.

▶ **Conjecture 2.** *There is an explicit construction of tournaments $T_1, T_2, \ldots$ such that $|V(T_n)| = \Theta(n)$ and $\vec{\omega}(T_n) = \Omega(n^{1/100})$.*

As we are unable to resolve Conjecture 1, we prove the following weaker result that gives some evidence that Conjecture 1 should be true. We turn our attention back to DIRECTEDCLIQUE problem, as it is much easier to construct directed graphs with high directed clique number. In particular, for the complete directed graph Q_n we have $\vec{\omega}(Q_n) = n$. The main result of this paper is the following.

▶ **Theorem 3.** *DIRECTEDCLIQUE is Σ_2^{P}-complete.*

Our feeling is that the reduction used to prove Theorem 3 is robust enough, so that a positive answer to Conjecture 2 should also give a positive answer to Conjecture 1 using similar techniques.

2 Main Result

The proof goes by a reduction from the following problem, which is a variation on the quantified boolean formula satisfaction problem. It is a natural Σ_2^{P}-complete problem with an easy reduction from EXISTENTIAL-2-LEVEL-SAT [12].

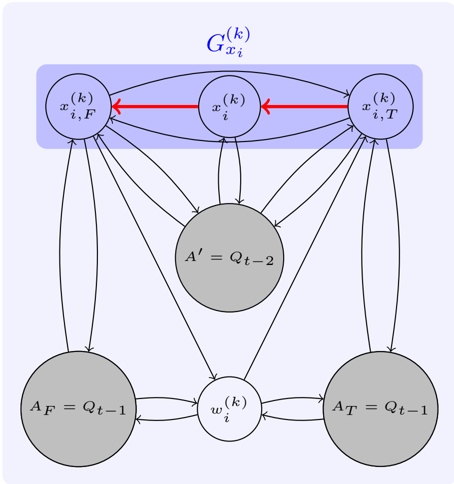

Figure 1 Binary gadget for occurrence of variable x_i in clause C_k.

Problem:	EXISTENTIAL-2-LEVEL-3-CNF
Input:	3-CNF formula $\varphi(x_1,\ldots,x_a,y_1,\ldots,y_b)$ with variables in two disjoint sets $\{x_1,\ldots,x_a\}$ and $\{y_1,\ldots,y_b\}$
Question:	Yes if and only if the following Boolean formula is true. $\exists_{x_1,x_2,\ldots,x_a} : \quad \neg\exists_{y_1,y_2,\ldots,y_b} : \quad \varphi(x_1,\ldots,x_a,y_1,\ldots,y_b)$

We are now ready to recall and prove the main theorem.

▶ **Theorem 3.** *DIRECTEDCLIQUE is Σ_2^{P}-complete.*

Proof. We present a reduction from EXISTENTIAL-2-LEVEL-3-CNF to DIRECTEDCLIQUE. Assume that we are given an instance $\langle X, Y, \mathcal{C}\rangle$ of EXISTENTIAL-2-LEVEL-3-CNF, where $X = \{x_1,\ldots,x_a\}$, $Y = \{y_1,\ldots,y_b\}$ are two disjoint sets of variables, and $\mathcal{C} = \{C_1,\ldots,C_c\}$ is a set of clauses with each clause having exactly three occurrences of three distinct variables in $X \cup Y$. For technical reasons, we assume that $c > 6$, as otherwise there are at most 18 variables and we can simply check all the possibilities. We set $t = 2c - 1$, and construct a graph G such that G admits a linear order $\ll$ of vertices such that $G^{\ll}$ does not contain K_{2c} as a subgraph if and only if the following Boolean formula

$$\exists_{x_1,x_2,\ldots,x_a} : \quad \neg\exists_{y_1,y_2,\ldots,y_b} : \quad C_1 \wedge C_2 \wedge \ldots \wedge C_c$$

is true. In the next paragraphs we describe gadgets necessary to construct the graph G.

Binary Gadget. For every occurrence of a variable $x_i \in X$ in a clause C_k we introduce a subgraph $G^{(k)}_{x_i}$, see Figure 1, consisting of the following:

1. Three disjoint complete directed graphs (gray nodes in Figure 1): A' with $t-2$ vertices, and A_F, A_T with $t-1$ vertices each.
2. Vertices $x^{(k)}_i$, $x^{(k)}_{i,T}$, $x^{(k)}_{i,F}$ and $w^{(k)}_i$.
3. Arcs as in Figure 1. An arc between a vertex and a gray node represents all possible arcs between the vertex and every vertex in the complete directed graph represented by the gray node.

Vertices $x_i^{(k)}$, $x_{i,T}^{(k)}$, and $x_{i,F}^{(k)}$ (in dark blue box in Figure 1) are later connected to other vertices in the graph. All the other vertices in the gadget, i.e. vertices in $A_F \cup A_T \cup A' \cup \left\{w_i^{(k)}\right\}$ are *internal* for the gadget and do not get any more adjacencies in the graph. The following claim captures the crucial property of the gadget.

▷ Claim 4. For any linear order $\ll$ on the vertices of $G_{x_i}^{(k)}$ such that the backedge graph $G_{x_i}^{(k)\ll}$ is K_{2c}-free we have that $\mathrm{BE}(G_{x_i}^{(k)}, \ll)$ contains exactly one of the edges $\{x_i^{(k)}, x_{i,F}^{(k)}\}$ or $\{x_i^{(k)}, x_{i,T}^{(k)}\}$ (arising from red arcs in Figure 1).

Proof. Let $\ll$ be any linear order on the vertices of $G_{x_i}^{(k)}$. Assuming the backedge graph is K_{2c}-free, we get $x_{i,F}^{(k)} \ll w_i^{(k)}$, as otherwise the $t-1$ vertices in A_F, $x_{i,F}^{(k)}$ and $w_i^{(k)}$ span a $2c$-clique in the backedge graph. Similarly, using the $t-1$ vertices in A_T, $x_{i,T}^{(k)}$, and $w_i^{(k)}$, we get $w_i^{(k)} \ll x_{i,T}^{(k)}$, and as a consequence $x_{i,F}^{(k)} \ll x_{i,T}^{(k)}$. If $x_{i,F}^{(k)} \ll x_i^{(k)} \ll x_{i,T}^{(k)}$ then these three vertices and the $t-2$ vertices in A' span a $2c$-clique in the backedge graph. Thus, exactly one of $x_i^{(k)} \ll x_{i,F}^{(k)}$ or $x_{i,T}^{(k)} \ll x_i^{(k)}$ holds and the claim follows. ◁

The idea behind the gadget is that the choice of the red edge remaining in the backedge graph corresponds to the valuation of variable x_i in clause C_k.

Copy Gadget. For every two occurrences of a single variable $x_i \in X$ in two different clauses C_k and C_ℓ we introduce a subgraph $G_{x_i}^{(k,\ell)}$, see Figure 2, consisting of the following:

1. Additional copy of a binary gadget (the middle layer in Figure 2).
2. Four disjoint complete directed graphs (gray nodes in Figure 1) A_1, A_2, A_3, A_4 with $t-3$ vertices each.
3. Arcs as in Figure 2.

All the new vertices introduced in the gadget are internal for the gadget and do not get any more adjacencies in the graph. Note that the gadget does not introduce any arcs between the vertices of $G_{x_i}^{(k)}$ and the vertices of $G_{x_i}^{(\ell)}$. The following claim captures the crucial property of the gadget.

▷ Claim 5. For any linear order $\ll$ on the vertices of $G_{x_i}^{(k)}$, $G_{x_i}^{(\ell)}$, and $G_{x_i}^{(k,\ell)}$ such that the backedge graph $G_{x_i}^{(k)\ll}$ is K_{2c}-free we have that $\mathrm{BE}(G_{x_i}^{(k,\ell)}, \ll)$ contains the edge $\{x_i^{(k)}, x_{i,F}^{(k)}\}$ if and only if it contains the edge $\{x_i^{(\ell)}, x_{i,F}^{(\ell)}\}$.

Proof. Assume that $\{x_i^{(k)}, x_{i,F}^{(k)}\}$ is an edge in $\mathrm{BE}(G_{x_i}^{(k,\ell)}, \ll)$. We get that $\{x_i^{(k,\ell)}, x_{i,T}^{(k,\ell)}\}$ is not an edge in $\mathrm{BE}(G_{x_i}^{(k,\ell)}, \ll)$, as otherwise the $t-3$ vertices in A_1 and the four surrounding vertices induce a $2c$-clique in the backedge graph. As the middle layer is a copy of the binary gadget, by Claim 4, we get that $\{x_i^{(k,\ell)}, x_{i,F}^{(k,\ell)}\}$ is an edge in $\mathrm{BE}(G_{x_i}^{(k,\ell)}, \ll)$. We get that $\{x_i^{(\ell)}, x_{i,T}^{(\ell)}\}$ is not an edge in $\mathrm{BE}(G_{x_i}^{(k,\ell)}, \ll)$, as otherwise the $t-3$ vertices in A_4 and the four surrounding vertices induce a $2c$-clique in the backedge graph. By Claim 4 applied to the binary gadget $G_{x_i}^{(\ell)}$, we get that $\{x_i^{(\ell)}, x_{i,F}^{(\ell)}\}$ is an edge in $\mathrm{BE}(G_{x_i}^{(k,\ell)}, \ll)$.

Similarly, if $\{x_i^{(k)}, x_{i,T}^{(k)}\}$ is an edge in $\mathrm{BE}(G_{x_i}^{(k,\ell)}, \ll)$, we get that $\{x_i^{(\ell)}, x_{i,T}^{(\ell)}\}$ is an edge in $\mathrm{BE}(G_{x_i}^{(k,\ell)}, \ll)$, and the claim follows. ◁

The idea behind the gadget is to ensure that red edges corresponding to all occurrences of a single variable in different clauses give a consistent valuation of variable x_i.

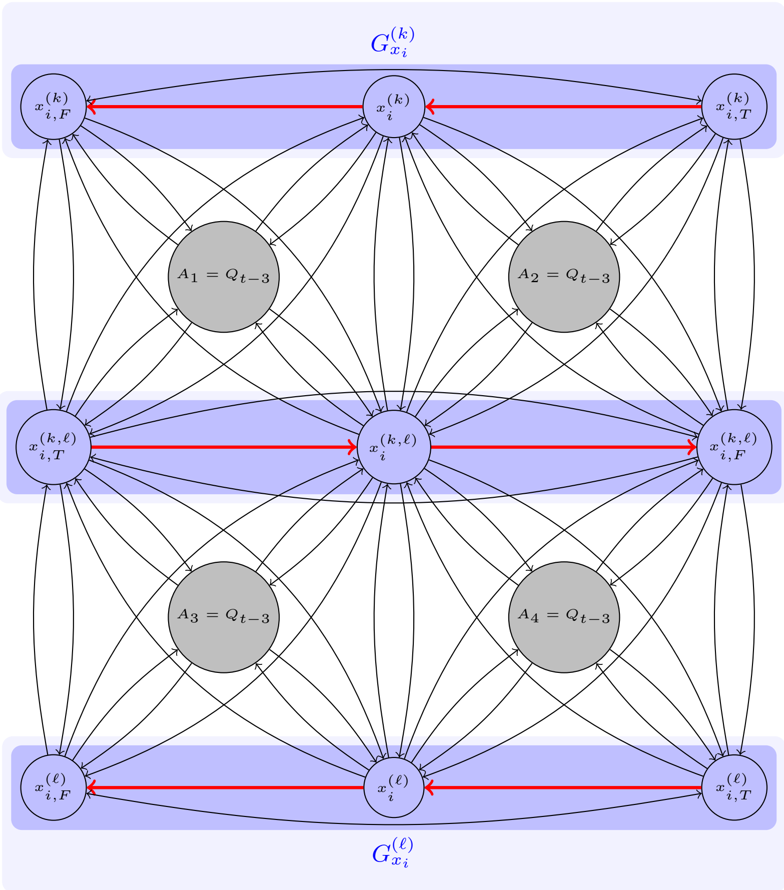

Figure 2 Copy gadget for occurrences of variable x_i in clauses C_k and C_ℓ. Blue boxes are copies of the binary gadget.

Clause gadget. For every occurrence of a variable $y_j \in Y$ in C_k, we introduce two vertices $y_{j,A}^{(k)}$, $y_{j,B}^{(k)}$ connected by two parallel arcs (in opposite directions). For every clause C_k we use 6 vertices in G in three groups: a pair of vertices for every occurrence z in C_k.

- If $z = y_j$ or $z = \neg y_j$ for some $y_j \in Y$, then the group consists of vertices $y_{j,A}^{(k)}$ and $y_{j,B}^{(k)}$.
- If $z = x_i$ for some $x_i \in X$, then the group consists of vertices $x_i^{(k)}$ and $x_{i,T}^{(k)}$ introduced in $G_{x_i}^{(k)}$.
- If $z = \neg x_i$ for some $x_i \in X$, then the group consists of vertices $x_i^{(k)}$ and $x_{i,F}^{(k)}$ introduced in $G_{x_i}^{(k)}$.

Reduction. We construct graph G by introducing a binary gadget for every occurrence of a variable in X, a copy gadget for every pair of occurrences of a single variable in X, and a clause gadget for every clauses in $\mathcal{C}$. Finally, for any two groups in two different clause gadgets if the groups correspond to different variables or correspond to the occurrences of the same variable with the same sign (both occurrences are positive or both are negative) we add all possible eight arcs between the two groups, see Figure 3 for an example. This concludes the description of the reduction. Note that the size of the constructed graph is polynomial in the size of the input formula and the reduction can be performed in polynomial time.

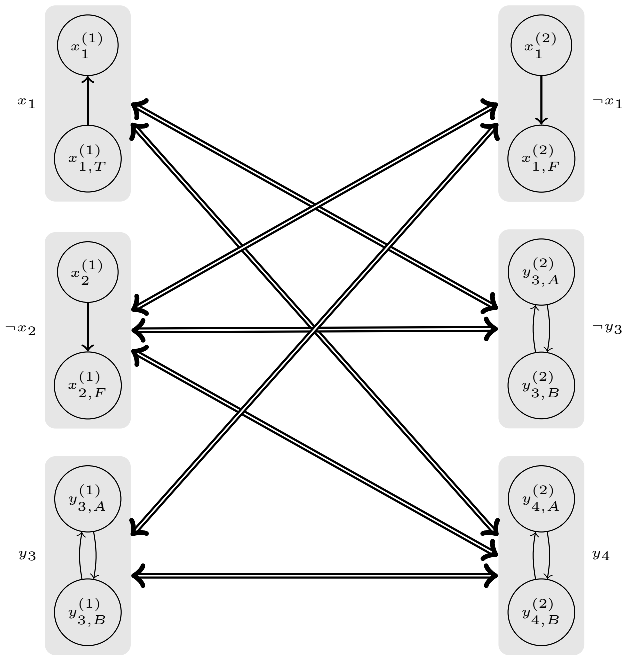

Figure 3 Clause gadgets for clauses $C_1 = x_1 \vee \neg x_2 \vee y_3$ (left) and $C_2 = \neg x_1 \vee \neg y_3 \vee y_4$ (right). Double edges represent all possible edges going from one group to another in any direction. There are no edges between vertices corresponding to x_1 in C_1 and in C_2, as these are negated occurrences. Similarly for occurrences of y_3.

Completeness. Assuming that $\langle X, Y, \mathcal{C}\rangle$ is a Yes instance, let $\nu = (\nu_1, \nu_2, \ldots, \nu_a)$ be a valuation of $x_1, x_2, \ldots, x_a$ such that the formula

$$\exists_{y_1,y_2,\ldots,y_b} : \quad (C_1 \wedge C_2 \wedge \ldots \wedge C_c)[x_1 = \nu_1, x_2 = \nu_2, \ldots, x_a = \nu_a]$$

is false. We construct a linear order $\ll$ such that the backedge graph $G^{\ll}$ does not contain a clique of size $2c$.

First, we use valuation ν to define the order in each binary gadget. Let $G^{(k)}_{x_i}$ be some binary gadget. If ν_i is true, we set $x_i^{(k)} \ll x_{i,F}^{(k)} \ll w_i^{(k)} \ll x_{i,T}^{(k)}$. Otherwise, we set $x_{i,F}^{(k)} \ll w_i^{(k)} \ll x_{i,T}^{(k)} \ll x_i^{(k)}$. We extend this order to the remaining internal vertices of the gadget arbitrarily, as the resulting backedge graph is the same for any resulting order. We show that the backedge graph induced by the vertices of the binary gadget is K_{2c}-free. As we have assumed $c > 6$, any clique with $2c$ vertices in the backedge graph must use at least one vertex in A_F, A_T or A'. As there are no arcs between these three sets, we can assume that the clique consists of all vertices in exactly one of the sets A_F, A_t or A' and possibly some of the remaing four vertices. Now, a simple case analysis similar to the proof of Claim 4 shows that the backedge graph induced by the vertices of the binary gadget is K_{2c}-free. If ν_i is true, the edge $\{x_i^{(k)}, x_{i,T}^{(k)}\}$ is in the resulting backedge graph. Otherwise, the edge $\{x_i^{(k)}, x_{i,F}^{(k)}\}$ is in the backedge graph. This way, the choice of the red edge in the backedge graph corresponds to the valuation ν_i of variable x_i. Observe that as the internal vertices of the gadget are not connected to any other vertices in the graph, no matter how we set the order $\ll$ on the remaining vertices, none of the internal vertices can be an element of K_{2c} in the backedge graph of the whole graph.

Similarly, we use valuation ν to define the order in each copy gadget. Let $G^{(k,\ell)}_{x_i}$ be some copy gadget. If ν_i is true, we set $x_i^{(k,\ell)} \ll x_{i,F}^{(k,\ell)} \ll w_i^{(k,\ell)} \ll x_{i,T}^{(k,\ell)}$. Otherwise, we set $x_{i,F}^{(k,\ell)} \ll w_i^{(k,\ell)} \ll x_{i,T}^{(k,\ell)} \ll x_i^{(k,\ell)}$. We extend this order to the remaining internal vertices of the gadget arbitrarily, as the resulting backedge graph is the same for any resulting order. Observe that when occurrence of x_i is of the same sign in clause C_k and clause C_ℓ, then the reduction adds eight additional arcs between vertices of $G^{(k)}_{x_i}$ and vertices of $G^{(\ell)}_{x_i}$ that are not present in Figure 2. Even with these additional edges, we show that the backedge graph induced by the vertices of the binary gadget is K_{2c}-free. As we have assumed $c > 6$, any clique with $2c$ vertices in the backedge graph must use at least one vertex in A_1, A_2, A_3 or A_4. As there are no arcs between these four sets, we can assume that the clique consists of all vertices in exactly one of the sets A_1, A_2, A_3 or A_4 and possibly some of the remaing vertices. Now, a simple case analysis similar to the proof of Claim 5 shows that the backedge graph induced by the vertices of the copy gadget is K_{2c}-free. As in a binary gadget, none of the internal vertices can be an element of K_{2c} in the backedge graph of the whole graph.

We combine the orders defined for all binary and all copy gadgets arbitrarily to get a linear order on all vertices introduced in these gadgets. Finally, we insert the vertices introduced in clause gadgets for occurrences of variables in Y into order $\ll$ arbitrarily. This concludes the construction of the linear order $\ll$.

We now have to show that the backedge graph $G^{\ll}$ does not contain a clique of size $2c$. Assuming to the contrary let X be a set of vertices that induce a clique of size $2c$ in the backedge graph. As observed before, none of the internal vertices is in X. In each binary gadget $G^{(k)}_{x_i}$, there are three non-internal vertices: $x_i^{(k)}$, $x_{i,T}^{(k)}$, and $x_{i,F}^{(k)}$. The definition of $\ll$ implies that either edge $\{x_i^{(k)}, x_{i,T}^{(k)}\}$ is in the backedge graph, or edge $\{x_i^{(k)}, x_{i,F}^{(k)}\}$ is in the backedge graph, but not both. Thus, there are at most two vertices in X in each binary gadget. Similarly, there are at most two vertices in X in each clause gadget, as the vertices in different groups in a single clause gadget are not connected. As $|X| = 2c$, we get that there are exactly two vertices in X in each clause gadget and the two vertices in a single clause gadget are in a single group that corresponds to some occurrence z in the clause C_k. If z is an occurrence of some $x_i \in X$ then we must have that the corresponding red edge is left in the backedge graph and we get that ν_i satisfies clause C_k. Now, we construct a valuation $\mu = (\mu_1, \mu_2, \ldots, \mu_b)$ of variables $y_1, y_2, \ldots, y_b$ as follows. If z is an occurrence of some $y_j \in Y$ then we set μ_j so that it satisfies clause C_k. As there are no edges between groups corresponding to opposite occurrences of a single variable y_j in different clause gadgets, this valuation of variables in Y is consistent. If valuation μ for some variable in Y is not decided yet, we set it arbitrarily. We get that the resulting valuation μ together with ν satisfies every clause in C_k. A contradiction.

Soundness. Assuming that $\langle G, t\rangle$ is a Yes instance, let $\ll$ be a linear order on vertices of G such that in $G^{\ll}$ there is no clique of size $2c$. Claim 4 and Claim 5 together imply that we can construct a valuation ν of the variables in X in the following way. We set ν_i to true if the edge $\{x_i^{(k)}, x_{i,T}^{(k)}\}$ is in the backedge graph for some occurrence of x_i in some clause C_k. By Claim 5, we have that the edge $\{x_i^{(\ell)}, x_{i,T}^{(\ell)}\}$ is in the backedge graph for every occurrence of x_i in any clause C_ℓ. Otherwise, we set ν_i to false.

We now have to show that the formula

$$\exists_{y_1, y_2, \ldots, y_b} : \quad (C_1 \wedge C_2 \wedge \ldots \wedge C_c)[x_1 = \nu_1, x_2 = \nu_2, \ldots, x_a = \nu_a]$$

is false. Assume to the contrary that there is a valuation $\mu = (\mu_1, \mu_2, \ldots, \mu_b)$ of $y_1, y_2, \ldots, y_b$ that together with ν satisfies every clause in the formula. We will use this valuation to construct a clique of size $2c$ in the backedge graph $G^{\ll}$. For every clause C_k we pick two vertices that form a group that corresponds to some variable that satisfies C_k. The construction of the reduction guarantees that there are all four possible edges between any two such groups in different clause gadgets. If the selected group corresponds to a variable in Y, then the vertices in the group are connected by an edge in the backedge graph. If the selected group corresponds to a variable in X, then by the definition of valuation ν the corresponding red edge is in the backedge graph. Thus, the selected vertices induce a clique of size $2c$ in the backedge graph. A contradiction. ◀

3 Upper bound

In this section we provide a simple upper bound on the directed clique number of a tournament.

▶ **Theorem 6.** *For every tournament T with $\overrightarrow{\omega}(T) \geq t \geq 1$, we have $|V(T)| \geqslant \binom{t+1}{2}$.*

Proof. The proof goes by induction on t. For the base case $t = 1$, we obviously have that the tournament needs to have at least $1 = \binom{1+1}{2}$ vertex. Now, for the induction step, let $t \geq 2$, and assume that the statement of the theorem holds for $t - 1$.

Let T be a tournament with $\overrightarrow{\omega}(T) \geq t$. As $\overrightarrow{\omega}(T) \geq t$, then for every linear order of $V(T)$, the backedge graph contains a clique K_t. Thus, there is at least one subset $U \subseteq V(G)$ of vertices that induces a copy of a transitive tournament TT_t. Let $U' = V(G) \setminus U$ be the complement of U, $T' = T[U']$ be the tournament induced by U', $t' = \overrightarrow{\omega}(T')$ be the directed clique number of T', and $\ll'$ be some linear order of U' with $\omega(T'^{\ll'}) = t'$.

As U induces a transitive subtournament in T, we can define order $\ll''$ on U such that for any $x, y \in U$ we have $x \ll'' y$ if and only if $(x, y) \in A(T)$. Consider a linear order $\ll$ on $V(G)$ that first puts all vertices in U' in order $\ll'$ and then puts all the vertices in U in order $\ll''$. As $\overrightarrow{\omega}(T) \geq t$, there is at least one copy of a clique K_t in the backedge graph $T^{\ll}$. We have that there are no edges between vertices in U in the backedge graph $T^{\ll}$. Thus, any clique in the backedge graph $T^{\ll}$ can include at most one vertex in U, and we have $t' \geq t - 1$. By the induction hypothesis, we get $|U'| \geq \binom{t}{2}$, and

$$|V(G)| = |U'| + |U| \geq \binom{t}{2} + t = \binom{t+1}{2}.$$ ◀

▶ **Corollary 7.** *For every tournament T, we have $\overrightarrow{\omega}(T) \leq \sqrt{2\,|V(T)|}$.*

Proof. Assuming to the contrary $\overrightarrow{\omega}(T) > \sqrt{2\,|V(T)|}$, we have $\binom{\overrightarrow{\omega}(T)+1}{2} > |V(T)|$, and a contradiction with Theorem 6. ◀

References

1 Pierre Aboulker, Guillaume Aubian, Pierre Charbit, and Raul Lopes. Clique number of tournaments, 2023. `doi:10.48550/arXiv.2310.04265`.

2 Pierre Aboulker, Nacim Oijid, Robin Petit, Mathis Rocton, and Christopher-Lloyd Simon. Computing the degreewidth of a digraph is hard, 2024. `arXiv:2407.19270`.

3 Sanjeev Arora and Boaz Barak. *Computational Complexity: A Modern Approach*. Cambridge University Press, 2009.

4 Guillaume Aubian. Computing the clique number of tournaments, 2024. `doi:10.48550/arXiv.2401.07776`.

5 Michael R. Garey and David S. Johnson. *Computers and Intractability: A Guide to the Theory of NP-Completeness*. W. H. Freeman, 1979.

6 Ilhee Kim. *On Containment Relations in Directed Graphs*. PhD thesis, Princeton University, 2013. URL: `http://arks.princeton.edu/ark:/88435/dsp012801pg46x`.

7 Víctor Neumann-Lara. The dichromatic number of a digraph. *Journal of Combinatorial Theory, Series B*, 33(3):265–270, 1982. doi:10.1016/0095-8956(82)90046-6.

8 Tung Nguyen, Alex Scott, and Paul Seymour. Some results and problems on tournament structure. *Journal of Combinatorial Theory, Series B*, 173:146–183, 2025. doi:10.1016/j.jctb.2025.02.002.

9 Marcus Schaefer and Christopher Umans. Completeness in the polynomial-time hierarchy: Part I: A compendium. *SIGACT News*, 33(3):32–49, 2002. doi:10.1145/582475.582484.

10 Marcus Schaefer and Christopher Umans. Completeness in the polynomial-time hierarchy: Part II. *SIGACT News*, 33(4):22–36, 2002. doi:10.1145/601819.601829.

11 Marcus Schaefer and Christopher Umans. Completeness in the polynomial-time hierarchy: A compendium, 2008. URL: https://ovid.cs.depaul.edu/documents/phcom.pdf.

12 Larry J. Stockmeyer. The polynomial-time hierarchy. *Theoretical Computer Science*, 3(1):1–22, 1976. doi:10.1016/0304-3975(76)90061-X.

13 Celia Wrathall. Complete sets and the polynomial-time hierarchy. *Theoretical Computer Science*, 3(1):23–33, 1976. doi:10.1016/0304-3975(76)90062-1.