Ann. Comb. 20 (2016) 139–176
DOI 10.1007/s00026-015-0295-4
Published online November 24, 2015

Annals of Combinatorics

Definability in the Substructure Ordering of Simple Graphs

Alexander Wires

School of Economic and Mathematics, Southwestern University of Finance and Economics, Chengdu 611130, P.R. China
slawkenbergius@hotmail.com

Received October 8, 2013

Mathematics Subject Classification: 05C99, 03C13, 06A06

Abstract. For simple graphs, we investigate and seek to characterize the properties first-order definable by the induced subgraph relation. Let $\mathcal{PG}$ denote the set of finite isomorphism types of simple graphs ordered by the induced subgraph relation. We prove this poset has only one non-identity automorphism **co**, and for each finite isomorphism type G, the set $\{G, G^{\mathbf{co}}\}$ is definable. Furthermore, we show first-order definability in $\mathcal{PG}$ captures, up to isomorphism, full second-order satisfiability among finite simple graphs. These results can be utilized to explore first-order definability in the closely associated lattice of universal classes. We show that for simple graphs, the lattice of universal classes has only one non-trivial automorphism, the set of finitely generated and finitely axiomatizable universal classes are separately definable, and each such universal subclass is definable up to the unique non-trivial automorphism.

Keywords: first-order definability, induced subgraphs, ordered set, universal classes

1. Introduction

By a simple graph, or just graph, we refer to a structure in a single binary relation which is symmetric and irreflexive. It is well known that interesting classes of graphs and graph properties can be described by omitting certain families of induced graphs; for example, disjoint sums of cliques avoid the path on three vertices, a graph is the line graph of a triangle-free graph if and only if it omits claws and diamonds ([3]), perfect graphs omit cycles of odd length greater than 3 and their complements [2], etc. Recently, it has been shown that those graphs omitting $K_{1,3}$ and its complement $K_{1,3}^{\mathbf{co}}$ play a role in the question of reconstructing graphs up to complementation ([13]). These results are often called forbidden graph characterizations.

For two finite graphs A and B, we write

$$A \leq B \text{ if and only if } A \text{ is isomorphic to an induced subgraph of } B.$$

As it appears in model theory, we will use the term substructure interchangeably with induced subgraph. Then a general forbidden graph characterization of a class of finite graphs $\mathcal{S}$ has the following form: a finite set of graphs $F_1, \ldots, F_n$ such that

$$G \in \mathcal{S} \text{ if and only if } \bigwedge_{i=1}^{n} G \not\geq F_i.$$

If we start with a set containing an isomorphic copy of every finite graph, then in the perspective of first-order logic, we can view $\mathcal{S}$ as a definable set in the relation $\leq$ using the finitely many parameters $F_1, \ldots, F_n$. In general, which properties of finite graphs can be captured using the substructure relation; more specifically, can we describe or characterize those sets of graphs which are definable using the full first-order language in the relation of substructure when applied to finite graphs? For example, forbidden graph characterizations correspond to the finite members of a class of graphs finitely axiomatized by universal sentences [14].

In this paper we will consider finite graphs in the isomorphic substructure relation $\leq$, and investigate definability in the full first-order language. Formally, we will take a set of finite graphs $\mathcal{G}$ whose vertices are over initial segments of positive integers. In this way, we have a complete set of representatives in which an isomorphic copy of every finite graph appears. The relation $\leq$ restricted to the graphs in $\mathcal{G}$ forms a quasi-ordered set $\langle \mathcal{G}, \leq \rangle$. Notice that A is isomorphic to B if and only if $A \leq B \leq A$. If we take the quotient of $\langle \mathcal{G}, \leq \rangle$ by the equivalence determined by isomorphism, we arrive at the poset $\langle \mathcal{PG}, \leq \rangle$ of finite isomorphism types ordered by substructure. A finite isomorphism type in $\mathcal{PG}$ can be represented by $H/\approx$ where $H \in \mathcal{G}$.

We explore definability in the structure $\langle \mathcal{PG}, \leq \rangle$. We say an n-ary relation R is definable if there is a first-order formula $\phi(x_1, \ldots, x_n)$ in the language of $\leq$ such that for $G_1/\approx, \ldots, G_n/\approx \in \mathcal{PG}$,

$$(G_1/\approx, \ldots, G_n/\approx) \in R \text{ if and only if } \langle \mathcal{PG}, \leq \rangle \models \phi(G_1/\approx, \ldots, G_n/\approx).$$

It is easy to see that the map which takes a graph G to its edge-complement $G^{\mathbf{co}}$ respects the substructure relation $\leq$, and so complementation determines an automorphism of $\langle \mathcal{PG}, \leq \rangle$. Our analysis proceeds by establishing certain species and auxiliary constructions are definable subsets if we use the type of the path on three vertices as a parameter; for example, circuits (Proposition 3.2), trees (Proposition 3.5), connected graphs (Proposition 5.5), and disjoint unions (Proposition 7.7). We shall establish how to definably compare the cardinality of graphs (Proposition 8.1) in a uniform manner which will afford a definable interpretation of addition. This analysis culminates in Section 9 with the proof of Theorem 9.2.

We then proceed further, and are able to characterize the expressive power of the substructure relation as capturing full second-order properties of graphs in the finite; that is, a subset of $\mathcal{G}$ is definable in the substructure relation using the constant P_3 precisely when they are the isomorphic copies in $\mathcal{G}$ of the finite models of a second-order sentence in the language of graphs (see Theorem 12.4).

This argument follows closely the approach taken in [9], passing through definable relations of the closely related small category Cat$\mathcal{G}$ described in Section 12. In fact, the present work can be seen as extending to the unordered structure of graphs the research in the series of papers “Definability in Substructure Ordering I–IV”([8–11]). In that work, Ježek and McKenzie investigated the scope of definable relations in the substructure relation for certain universal classes of finite structures with an underlying partial order; namely, ordered sets, meet-semilattices, lattices, and distributive lattices. Their study addresses the general question of *positive definability* for the universal theories of these structures. With our Theorem 9.2, we will be able to conclude positive definability for simple graphs (see Theorem 10.3).

The discussion of positive definability and universal theories will be delayed until Section 10 after Theorem 9.2 is established.

For a highly readable introduction to the basic notions of structures, language, and first-order logic which we use, please consult [1, V. 1].

2. Preliminaries and Small Orders

Formally, a graph $G = \langle V(G), E(G)\rangle$ is a first-order structure with universe, or vertex set, V and a binary relation E, or edge relation, which is irreflexive and symmetric. We shall abuse notation slightly and write $v \in G$ to refer to a vertex of G. For two vertices $u, v \in G$, the edge relation for u and v is denoted as $u \sim v$ and it is said that u and v are adjacent. We do not consider the graph on an empty set of vertices.

In order to formally investigate definable relations in the isomorphic substructure relation among finite graphs, we need to consider the appropriate structure with universe a set. Let $\mathcal{G}$ denote the set of finite graphs with vertices over initial segments of positive integers. Then the structure $\langle \mathcal{G}, \leq\rangle$ is a quasi-ordered set in which an isomorphic copy of every finite graph appears. An n-ary relation R is definable in $\langle \mathcal{G}, \leq\rangle$ if there is a first-order order formula $\phi(x_1,\dots,x_n)$ in the fundamental relation $\leq$ such that for $G_1,\dots,G_n \in \mathcal{G}$,

$$(G_1,\dots,G_n) \in R \text{ if and only if } \langle \mathcal{G}, \leq\rangle \models \phi(G_1,\dots,G_n).$$

It is a standard result, and easy to prove by induction on the length of formulas, that definable relations are closed under automorphisms; in this case, if ϕ is an automorphism of $\langle \mathcal{G}, \leq\rangle$ and R an n-ary definable relation, then for all $G_1,\dots,G_n \in \mathcal{G}$, $(G_1,\dots,G_n) \in R$ if and only if $(\phi(G_1),\dots,\phi(G_n)) \in R$. There is an obvious automorphism of $\langle \mathcal{G}, \leq\rangle$ which is defined by edge complementation and denoted by **co**; that is, $\mathbf{co}(G)$ is the graph over the same set of vertices as G, but $u \sim v$ in $\mathbf{co}(G)$ if and only if $u \not\sim v$ in G. Unfortunately, $\langle \mathcal{G}, \leq\rangle$ has too many “inconsequential” automorphism since any map of $\mathcal{G}$ which arbitrarily permutes graphs within the isomorphism classes is an automorphism of $\langle \mathcal{G}, \leq\rangle$. If we take the quotient of $\langle \mathcal{G}, \leq\rangle$ by the natural equivalence relation determined by isomorphism, we arrive at the poset $\langle \mathcal{PG}, \leq\rangle$ of finite isomorphism types in which the definable relations are the quotients by isomorphism of the definable relations of $\langle \mathcal{G}, \leq\rangle$. Edge-complementation naturally induces an automorphism of $\langle \mathcal{PG}, \leq\rangle$ and we shall see in Section 9 that this is the only non-trivial automorphism.

While we are interested in definability in the poset $\langle \mathcal{PG}, \leq\rangle$, it will be more convenient to work within the closely related quasi-ordered set $\langle \mathcal{G}, \leq\rangle$ where we will speak of graphs definable up to isomorphism rather than definable isomorphism types. If

$G < H$, but there does not exist F such that $G < F < H$, then we write $G \prec H$ and say H covers G. It is easy to see that $G \prec H$ if and only if $G \leq H$ and $|H| = |G| + 1$. When $A \leq B$, we can then identify A with a particular induced subgraph U of B such that $U \approx A$. For example, if $A \prec B$, then we will say that B is formed from A by adding an additional vertex v to A and possibly some additional edges connecting v to vertices of A. If v is a vertex of G, then $G - v$ will denote the induced subgraph on the vertices of G omitting v; that is, the induced subgraph on the vertex set $V(G) \setminus \{v\}$.

It follows that the poset $\langle \mathcal{PG}, \leq \rangle$ is naturally graded according to cardinality, and so for each fixed positive integer n those isomorphism types at height n (having cardinality n) are definable as having a maximal n-element chain in their principal order ideals. Notice this definition requires a fixed n, and so those types with cardinality $n+1$ require a different package of formulas to define them. We shall see later in Section 8 how to capture cardinality in a uniform manner.

The *complete graph*, or *clique*, on m vertices is denoted by K_m and is characterized as having every possible edge. The *trivial* graph on m vertices is denoted by N_m and has no edges. The *path* on n vertices is denoted by P_n and is a graph isomorphic to the graph $v_1 \sim v_2 \sim \cdots \sim v_n$ with no additional edges other than the ones specified. The *circuit* (or *cycle*) C_n is formed from the path P_n by adding only one additional edge $v_n \sim v_1$. Our notion of a connected graph and connected component are standard.

For two graphs G and H, we form the disjoint sum $G + H$ by taking the disjoint union of the two sets of vertices and allowing only those edges coming from G and H. Clearly, if $A \approx G$ and $B \approx H$, then $A + B \approx G + H$. We may consider the sum of more than two graphs, and so when taking many factors $\{G_i\}$ we can write the sum as $\sum G_i$; of course, this yields a convenient general notation for graphs as the disjoint sum of their connected components.

Given two graphs G and H, the join $G \bigvee H$ is formed by taking $G + H$ and adding every possible edge of the form $u \sim v$ where $u \in G$ and $v \in H$; for example, $K_{p+q} \approx K_p \bigvee K_q$. Again, it is easy to see that if $A \approx G$ and $B \approx H$, then $A \bigvee B \approx G \bigvee H$.

The diamond $\mathbf{D}$ is formed from K_4 by removing a single edge.

The graph on four vertices with only the edges $u \sim v \sim x$ and $v \sim y$ will be denoted by $K_{1,3}$. This graph is often referred to as the *claw*.

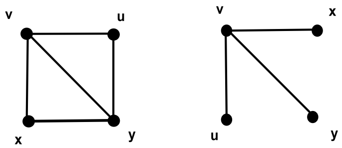

Figure 1: Graphs $\mathbf{D}$ and $K_{1,3}$.

At this point, we will add the constant P_3 representing a path on three vertices to form the pointed quasi-ordered structure $\langle \mathcal{G}, \leq, P_3 \rangle$. Definable relations in this structure are determined by first-order formulas in the language $\{\leq, P_3\}$ and are said to be definable with constants (or with parameters). Definable relations in the language $\{\leq\}$ are said to be definable without constants. Since $\langle \mathcal{PG}, \leq \rangle$ has the non-trivial automorphism **co**, we will be investigating relations definable with constants and so unless otherwise specified, definability will refer to the language $\{\leq, P_3\}$. The next lemma is one of the few examples where we make the distinction.

Lemma 2.1. *The set of complete or trivial graphs is definable without constants.*

Proof. A graph G is complete or trivial if and only if $G\downarrow$ is a chain.

To see this, suppose $G \not\approx K_n, N_k$ for any $n, k \geq 1$. Then $|G| \geq 3$ and so there exist vertices u, v, x, y such that $|\{u, v, x, y\}| \geq 3$ and $u \sim v$ and $x \not\sim y$. Then $K_2 \leq G$ and $N_2 \leq G$, but K_2 and N_2 are incomparable. ∎

For a circuit C_n we may construct a new graph $C_n \rightarrow_1$ by adding only one new edge $u \sim x$ where x is some new vertex and u is an arbitrary vertex of C_n. Different choices of u result in isomorphic graphs, and so the construction is well defined on isomorphism types. The same construction for K_n in place of C_n is denoted by $K_n \rightarrow_1$. The graph $K_3 \rightarrow_1$ is sometimes referred to as the *paw*.

Ultimately, we will show every finite graph is definable in Proposition 9.1, but as a first step we will capture those graphs on at most four vertices in $\langle \mathcal{G}, \leq, P_3 \rangle$.

Proposition 2.2. *Every graph on at most four vertices is definable.*

Proof. By Lemma 2.1, the set $\{K_3, N_3\}$ is definable since they are complete or trivial and of order 3. Then $\{P_3, K_2 + N_1\}$ is definable as the remaining graphs of order three; therefore, $K_2 + N_1$ is definable since we can call up the constant P_3.

Now P_4 is definable as the unique element covering P_3 and $K_2 + N_1$, but not covering anything in $\{K_3, N_3\}$.

The circuit C_4 is the element with the unique lower cover P_3.

We would like to define $P_3 + N_1$. Notice that $P_3 + N_1$ and $K_3 \rightarrow_1$ both cover P_3, $P_2 + N_1$, and some graph in $\{K_3, N_3\}$. We shall have recourse to uniquely define a graph which covers C_4 which will help separate $P_3 + N_1$ from $K_3 \rightarrow_1$. First, we see that $C_4 + N_1$ can be defined as the unique cover of C_4 which has exactly two subcovers A and B, and if $A \approx C_4$, then $B \not\approx P_4$ and covers both P_3 and $K_2 + N_1$. We can then recover $P_3 + N_1$ as the unique subcover of $C_4 + N_1$ which is not isomorphic to C_4.

N_3 is the unique subcover of $P_3 + N_1$ which is not P_3 nor $K_2 + N_1$

K_3 is then defined as the complete or empty graph of order three which is not N_3; consequently, K_4 is also definable. It follows that K_2 and N_2 are separately definable.

The graph $K_3 \rightarrow_1$ is definable as the unique element with lower covers P_3, $K_2 + N_1$, and K_3.

$K_3 + N_1$ is the only cover of $P_2 + N_1$ which also covers K_3 and is not isomorphic to $K_3 \rightarrow_1$.

$K_2 + K_2$ is the element with unique lower cover $P_2 + N_1$.

$K_2 + N_2$ is the unique element with only $P_2 + N_1$ and N_3 as lower covers.

The graph $\mathbf{D}$ is definable as the unique cover of P_3 not equal to $K_3 \rightarrow_1$, but which also covers K_3.

The graph $K_{1,3}$ is definable as the element with P_3 and N_3 as the only lower covers.

This accounts for every graph on at most four vertices. ∎

Since K_3 and N_3 are separately definable, Lemma 2.1 implies that the sets of complete and trivial graphs are separately definable. Here we point out that the first paragraph of the proof of Proposition 2.2 shows that the set $\{P_3, K_2 + N_1\}$ is definable without constants.

The clique number $\omega(H)$ is the order of the largest clique which embeds in G, and the independence number $\alpha(A)$ is the order of the largest trivial graph which embeds in G; consequently, both these parameters are definable. To be specific, both relations $\{(A,B)\colon A$ is complete, $\omega(B)=|A|\}$ and $\{(A,B)\colon A$ is trivial, $\alpha(B)=|A|\}$ are definable.

3. Circuits, Paths, and Trees

A graph is *acyclic* (or a *forest*) if it avoids circuits, and a *tree* if it is a connected forest. Our first task then is to show how to capture circuits.

Lemma 3.1. $\{A\colon A\approx\sum C,\ C$ *is a circuit* $, |C|>3\}$ *is definable.*

Proof. The claim is that A is in this set if and only if $|A|\geq 4$, and $P_3\leq A$, $K_{1,3}\nleq A$, $K_3\nleq A$, and A has a unique lower cover.

It is straightforward to see that these conditions are necessary. We must show that they are sufficient.

Suppose A is a graph which satisfies the conditions. If A is a circuit, then we are done. Assume A is not a circuit. Note that $P_3\leq A$ implies A is not trivial.

Claim. A is the disjoint sum of circuits and paths.

Proof. First, note the maximum degree of every vertex in A is two. Suppose not. Let v be a vertex of A such that its neighborhood $N(v)=\{u: v\sim v\}$ has at least three vertices. If any two vertices of $N(v)$ are adjacent, then $K_3\leq A$, a contradiction. Since $|N(v)|\geq 3$ and no vertices are adjacent, the induced subgraph on the vertices $N(v)\cup\{v\}$ embeds a copy of $K_{1,3}$, another contradiction.

Since every vertex has maximum degree two, it is not too hard to see that A must be the disjoint sum of circuits and paths. This finishes the claim.

Suppose A is the disjoint sum of k circuits and r paths. Let v be a vertex of some circuit and u a vertex of some path. Then $A-v$ has $k-1$ circuits, but $A-u$ still has k circuits. This contradicts the fact that A has a unique lower cover. Suppose A is only the disjoint sum of t paths. Since $P_3\leq A$, some path P in the sum of A has at least three vertices. Let x be a terminal vertex in P and let y be a vertex of degree two in P. Then $A-x$ is the disjoint sum of t paths, but $A-y$ is the disjoint sum of $t+1$ paths; a contradiction.

It must be the case that A is a disjoint sum of circuits. Since A has a unique lower cover, all the circuits must have the same length. ∎

For any $n\neq m$, we see that C_n and C_m are incomparable, and therefore, any two disjoint sums in the previous relation are comparable if and only if they are disjoint sums over the same isomorphic circuit. In this case, the set of such sums over the same circuit are naturally linearly ordered according to the number of components; of course, the minimal elements are just the circuits.

Proposition 3.2. *The set of circuits is definable.*

Proof. If we denote the definable set in Lemma 3.1 as *CSUM*, then we have the following definable relation $R = \{(A,B) \colon A,B \in CSUM, A \leq B\}$. It follows from the above discussion that C is a circuit if and only if

$C \approx K_3$, or

$|C| > 3$, $C \in CSUM$, and $\forall B[(B,C) \in R \rightarrow B \approx C]$. ∎

Corollary 3.3. *The following sets are definable*:

(1) *forests,*

(2) *paths,*

(3) *disjoint sum of paths.*

Proof. It is easy to see that a graph is strictly below some circuit if and only if it is a disjoint sum of paths. Paths are just those elements which are the unique lower covers of a circuit. ∎

Recall the following definition of distance.

Definition 3.4. *For a,b in the same connected component of Γ, let $d(a,b)$ equal one less than the cardinality of the shortest path in Γ connecting a to b. The diameter of Γ is then taken to be*

$$d(\Gamma) = \max\{d(a,b) \colon a,b \text{ are in the same connected component of } \Gamma\}.$$

When a,b are in different connected components, set $d(a,b) = \infty$.

The distance between any two vertices a and b in the same connected component is always realized by some path, say, $P = a \sim x_1 \sim \cdots \sim x_n \sim b$. If $x_i \sim x_j$ for some $j > i+1$, then we may construct a shorter path from a to b, contradicting the minimality of P; thus, the distance is always realized by an embedded path $P \leq \Gamma$. Since Γ is finite, the diameter is always realized by some path, and thus, by an embedded path. While it is not always true that $A \prec B$ implies $d(A) \leq d(B)$ — consider $A \approx P_4$ and $B \approx C_5$ — it is true for the class of forests. It is also the case that the diameter of a forest is just the length of the largest induced subpath.

Proposition 3.5. *The set of trees is definable.*

Proof. The claim is that T is a tree if and only if T is a forest and for any forest D such that $T \prec D$ we have $d(D) \leq d(T)+1$.

Suppose T is a tree, then whenever D is a forest such that $T \prec D$, we can construct D from T by adding a new vertex x and at most a single new edge $u \sim x$ where $u \in T$. Let $P \leq D$ realize the diameter of D. If $x \notin P$, then $d(T) \geq d(D)$. If $x \in P$, then x is a terminal vertex in the path P since it has degree one; therefore, we have a path $\overline{P} \leq T$ such that P is equal to adjoining x to the end of $\overline{P} \leq T$. Then, $|\overline{P}| \leq d(T)$, which implies $d(D) = |P| \leq d(T)+1$.

Conversely, if F satisfies the conditions, then we may write $F = \sum_{i=1}^{m} F_i$ where each F_i is a tree. Note there exists $1 \leq k \leq m$ such that $d(F) = d(F_k)$. If $m > 1$, then choose $j \neq k$ and construct R, a cover for F, in the following manner: take F and a new vertex $v \notin F$, and add two new edges $a \sim v$ and $b \sim v$ where $a \in F_j$ and $b \in F_k$ such that b is an end-vertex of a path $P \leq F_k$ which realizes $d(F_k)$. It is easy to see that $d(R) > d(F_k)+1$; thus, we must have $m = 1$, which implies F is a connected forest. ∎

The above argument utilized the fact that the diameter for acyclic graphs was definable. If in general, the diameter of a graph was a definable property, then one would hope an argument similar to that of Proposition 3.5 would yield the definability of the set of connected graphs. Explicitly, one would need the result that Γ is connected if and only if every upper cover increases the diameter by at most one. Unfortunately, this is not true since one can find counterexamples among trees and their covers which are not forests. We will have to take a different approach to capture connected graphs in Section 5.

We record the following corollary for use in the next section.

Corollary 3.6. *$\{(C,\Gamma)\colon C$ is circuit and $\Gamma \approx C+C\}$ is definable.*

Proof. (C,Γ) is in the relation if and only if C is a circuit and

$C \approx K_3$ and Γ is a disjoint sum of cliques, $\omega(\Gamma) = 3$ and $\alpha(\Gamma) = 2$, or

$C \approx C_4$ and C_4 is the unique circuit strictly below Γ, $K_{1,3} \not\leq \Gamma$, Γ has a unique lower cover, and $\alpha(\Gamma) = 4$, or

$|C| > 4$, $\Gamma \in CSUM$, $C_m < \Gamma$, and there does not exist $R \in CSUM$ such that $C_m < R < \Gamma$.

The case where $C \approx C_4$ requires some explanation. The first two conditions imply that Γ is a disjoint sum of copies of C_4 and possibly of some paths. Since Γ has a unique lower cover, there cannot be any paths present in the disjoint sum. The condition $\alpha(\Gamma) = 4$ implies that there are only two copies of C_4 in the sum, that is, $\Gamma \approx C_4 + C_4$. ∎

4. Addition with Paths and Circuits

We would like to accomplish addition in the cardinality of paths and circuits. Observe that the lower covers of a path P_k are precisely the path P_{k-1} and the disjoint sums $P_r + P_t$ where $r+t = k-1$.

Lemma 4.1. *$\{(P,G)\colon P \approx P_m, G \approx P_m + P_m\}$ is definable.*

Proof. The claim is that $G \approx P_m + P_m$ if and only if

$m = 1$ and $G \approx N_2$, or

$m \geq 2$ and $G \prec E \prec C_{m+1} + C_{m+1}$ for some E such that G is acyclic.

To see this one merely has to observe that any acyclic G such that $G \prec E \prec C_{m+1} + C_{m+1}$ must come from deleting a single vertex from each of the components in the sum. That $C_{m+1} + C_{m+1}$ is definable is precisely Corollary 3.6. ∎

Proposition 4.2. *$\{(A,B,P)\colon A \approx P_n, B \approx P_m, P \approx P_{n+m}\}$ is definable.*

Proof. The claim is that $P \approx P_{n+m}$ if and only if P is a path, there exists a path R such that $P \prec R$, and there exists $G \prec R$ such that

(1) G is not a path,
(2) $P_n \leq G$ and $P_m \leq G$,
(3) if Q is a path such that $Q \leq G$, then $Q \leq P_n$ or $Q \leq P_m$
(4) if $P_m \leq P_n$ we have $P_m + P_m \leq G$, and if $P_n \leq P_m$ we have $P_n + P_n \leq G$,

and for any path E such that $F \prec E$ and F satisfies (1)–(4), then $R \leq E$.

To show these conditions are sufficient, suppose P satisfies the conditions and is covered by a path R and $G \prec R$. Assume $R \approx P_s$ and so, $G \approx P_r + P_t$ where $r+t = s-1$. Without loss of generality, we may take $n \geq m$. If $r > n$ or $t > n$, then $P_{n+1} \leq G$, but $P_{n+1} \nleq P_n$ and $P_{n+1} \nleq P_m$ which contradicts (3); thus, $r, t \leq n$. By (2), $r = n$ or $t = n$, and so we may assume $r = n$. By (4), $t \geq m$ and so we can conclude that $n+m \leq s-1 \leq n+n$. For $f-1$ in the interval $[n+m, n+n]$, each P_f has a lower cover which satisfies (1)–(4), and so we must have $s-1 = n+m$ by the requirement of minimality. This implies $P \approx P_{n+m}$.

Clearly, $P_n + P_m \prec P_{n+m+1}$ satisfies (1)–(4), and by the above argument any disjoint sum of two paths which satisfies (1)–(4) must be covered by a path P_s with $n+m+1 \leq s \leq n+n+1$ where $n \geq m$. This establishes that the conditions are necessary, and completes the proof of the proposition. ∎

As a corollary we may establish the definability of the disjoint sum of two paths.

Corollary 4.3. $\{(A,B,P)\colon A,B$ *are paths and* $P \approx A+B\}$ *is definable.*

Proof. $P \approx P_n + P_m$ if and only if $P \prec P_{n+m+1}$, and

(1) $P_n \leq P$ and $P_m \leq P$,
(2) if Q is a path such that $Q \leq P$, then $Q \leq P_n$ or $Q \leq P_m$.

If P satisfies the conditions, then $P \approx P_{n+m}$ or $P \approx P_r + P_t$ where $r+t = n+m$. Without loss of generality, assume $n \geq m$. Since $P_{n+1} \leq P_{n+m}$, by (2) we see that $P \approx P_r + P_t$. Condition (1) implies $r \geq n$ or $t \geq n$. If $r > n$, or $t > n$, then $P_{n+1} \leq P$ and we arrive at a contradiction of (2); thus, $r = n$ or $t = n$ which implies $t = m$ or $r = m$, respectively.

As the necessity of the conditions is immediate, we have established the result. ∎

Since paths are the unique lower covers of circuits, we can also accomplish addition with the definable set of circuits in the obvious way.

Corollary 4.4. $\{(A,B,E)\colon A \approx C_n, B \approx C_m, E \approx C_{n+m}\}$ *is a definable relation.*

Corollary 4.5. $\{(A,B,E)\colon A \approx N_p, B \approx N_q, E \approx N_{p+q}\}$ *is a definable relation.*

Proof. Use the fact $\alpha(P_m) = \lfloor m/2 \rfloor$, and that whenever P, Q are distinct paths with $\alpha(P) = \alpha(Q)$, then the order of Q is even precisely when $P \prec Q$. The result now follows from Corollary 4.2. ∎

5. Connectedness

In this section, we will show that the set of connected graphs is definable.

Lemma 5.1. $\{(C,E)\colon C \approx C_m$ *and* $E \approx C_m + N_1\}$ *is definable.*

Proof. The claim is that $E \approx C_m + N_1$ if and only if

$m = 3$ and $E \approx K_3 + N_1$, or
$m > 3$ and $C_m \prec E$, $K_{1,3} \nleq E$, $K_3 \nleq E$.

The necessity of the conditions is immediate.

For sufficiency, suppose E satisfies the conditions and $E \not\approx K_3 + N_1$. Then we may form E from C_m by adding an additional vertex v and possibly some new edges connecting C_m to v. Suppose there exist $a, b \in C_m$ such that $a \sim v$ and $b \sim v$. If $a \sim b$, then the induced subgraph on the vertices $\{a, b, v\}$ is isomorphic to K_3, a contradiction. It must be the case that $a \not\sim b$, but then $K_{1,3} \leq E$ when we consider the induced subgraph on the vertices $\{z, a, w, v\}$ where $z \sim a \sim w$ and $z, w \in C_m$, another contradiction. So there can be at most one new edge. Since $m > 3$, we see that $K_{1,3} \leq E$ if there is just one additional edge; therefore, $E \approx C_m + N_1$. ∎

Let $Path_{\geq 2}$ denote the set of graphs which are disjoint sums of paths with no isolated vertices.

Lemma 5.2. *$Path_{\geq 2}$ is definable.*

Proof. If we can show that the set of graphs which are disjoint sums of paths with isolated vertices is definable, then the lemma will follow. Recall that the set of circuits forms an antichain under the substructure ordering; however, their unique subcovers, the set of paths, are linearly well-ordered. This implies there is a first-order definable well-ordering $\leq_*$ on circuits defined by

$$C \leq_* D \text{ if and only if } P \leq Q,$$

for circuits C and D and $P \prec C$ and $Q \prec D$.

The claim is that G is a disjoint sum of paths with isolated vertices if and only if

$|G| = 1$ and $G \approx N_1$, or
$|G| = 2$ and $G \approx N_2$, or
$|G| = 3$ and $G \approx N_3$, or $G \approx K_2 + N_1$, or
$|G| > 3$ and

(1) G is a disjoint sum of paths,
(2) if C is a circuit such that $G \leq C$, and C is the smallest circuit D under $\leq_*$ such that $G \leq D$, then there exist circuits E and F such that $E \prec_* F \prec_* C$ and $G \leq E + N_1$.

The preceding observations and Corollary 3.3 guarantee that these conditions are definable. For necessity, assume G is a disjoint sum of paths with isolated vertices and write $G \approx N_1 + \sum_{i=1}^{r} P_i$. If $n = \sum_{i=1}^{r} |P_i|$, then C_{n+r+2} is the circuit of smallest cardinality which embeds G. Then $\sum_{i=1}^{r} P_i \leq C_{n+r}$ and we set $E \approx C_{n+r}$ for condition (2).

For sufficiency, assume G satisfies the conditions and that $|G| > 3$. By (1), G is a disjoint sum of paths. It is easy to see that if G has no isolated vertices, then $G \leq C_k$ exactly when $G \leq C_k + N_1$. Let C be the smallest circuit under $\leq_*$ such that $G \leq C$. If G has no isolated vertices then using (2), $G \leq E$ where E is a circuit $E <_* C$; a contradiction. It must be the case that G has isolated vertices. ∎

Proposition 5.3. *$\{(X, N, G) \colon N \approx N_m$ and $G \approx X + N_m\}$ is definable.*

Proof. The claim is that $G \approx X + N_m$ if and only if

G is trivial, X is trivial and $\alpha(G) = \alpha(X) + m$, or
G is not trivial, and

(1) X is not trivial and $X < G$,
(2) $\alpha(G) = \alpha(X) + m$,
(3) for every circuit $\overline{C}$ there exists a circuit C such that
 (a) $\overline{C} <_* C$,
 (b) there exists Γ such that
  (i) $X \leq \Gamma$ and $C \leq \Gamma$, and for all $H \leq \Gamma$, $X \leq H$ and $C \leq H$ imply that $H \approx \Gamma$,
  (ii) $G \leq \Gamma$,
  (iii) for all $R \leq \Gamma$ such that $C \prec R$, $C + N_1 \approx R$,
(4) for all $B \in \mathit{Path}_{\geq 2}$, $B \leq G$ implies $B \leq X$.

We first tackle the argument for sufficiency. Suppose G satisfies the conditions. We may assume G is not trivial; otherwise, definability follows from addition with trivial graphs provided by Corollary 4.5. We can represent G as $G \approx E + P$ where E is the disjoint sum of connected components which are not paths and P is the disjoint sum of all the connected components which are paths. Also, write $X \approx A + Q$ where A the disjoint sum of connected components which are not paths and Q is the disjoint sum of all the connected components which are paths. Suppose $n = |Q|$ and Q has r components (all of which are paths). Take $C_{n+r} = \overline{C}$ in (3). Let C be the circuit with $C_{n+r} <_* C$ given by (3a).

Let Γ be the graph whose existence is guaranteed by (3b). Condition (iii) implies that any copy of the circuit C in Γ must appear as a connected component, and since $C \leq \Gamma$, we can write $\Gamma \approx C + K$ for some sum of connected graphs K. Since $X \leq \Gamma$, we must have $A \leq K$. Because C^* was chosen large enough such that $Q \leq C$, we have $X \leq C + A \leq C + K \approx \Gamma$; thus, by (i) we have $\Gamma \approx A + C$. Note that

$$A + Q \approx X < G \approx E + P \leq \Gamma \approx A + C$$

implies $E \approx A$ and so, $G \approx A + P$ where $Q \leq P \leq C$.

We can further write $Q \approx F + N_t$ and $P \approx H + N_r$ where $F, H \in \mathit{Path}_{\geq 2}$. It is easy to see that whenever K is maximal among those graphs $\Phi \in \mathit{Path}_{\geq 2}$ such that $\Phi \leq X \approx A + Q \approx A + F + N_t$, then $K \approx J + F$ where $J \in \mathit{Path}_{\geq 2}$ and is maximal for $J \leq A$. Take $\bar{J} \in \mathit{Path}_{\geq 2}$ such that $\bar{J} \leq A$ and is of maximum cardinality. The condition $Q \leq P$ implies $F \leq H$. If $F < H$, then $\bar{J} + H \in \mathit{Path}_{\geq 2}$ and $\bar{J} + H \leq G$, and so by (4) we must have $\bar{J} + H \leq A + F$ which contradicts the choice of $\bar{J}$. It must be the case that $F \approx H$. Condition (2) implies $N_r \approx N_t + N_m$ and so,

$$G \approx A + P \approx A + H + N_r \approx A + F + N_t + N_m \approx A + Q + N_m \approx X + N_m.$$

To prove these conditions are necessary, assume X is not trivial and write $X \approx A + Q$ as before with $n = |Q|$ and r such that Q has r components (all of which are paths). Then for any $C_{n+r+2m} <_* C$ notice that $Q \leq C$. We may then take $\Gamma \approx A + C$ and it is straightforward to check conditions (3) and (4) are satisfied. ∎

Proposition 5.3 actually yields more than is explicitly stated. What we have shown is that there is a first-order formula $\Psi(x, y, z, w)$ in the language of $\langle \mathcal{G}, \leq \rangle$ such that $\langle \mathcal{G}, \leq \rangle \vDash \Psi(A, N, G, P_3)$ if and only if N is trivial and $G \approx A + N$. If we apply the complementation automorphism, we see that $\langle \mathcal{G}, \leq \rangle \vDash \Psi(B, K, H, \mathbf{co}(P_3))$ if and only if $\mathbf{co}(N) = K$ is complete and $H \approx \mathbf{co}(A+N) = \mathbf{co}(A) \bigvee \mathbf{co}(N) = B \bigvee K$. Since $\mathbf{co}(P_3) = K_2 + N_1$ is definable in $\langle \mathcal{G}, \leq, P_3 \rangle$, there is a first-order formula $\gamma(x, w)$ such that $\langle \mathcal{G}, \leq \rangle \vDash \gamma(A, P_3)$ if and only if $A \approx K_2 + N_1$. We can then take the formula

$$\exists w \Psi(x, y, z, w) \wedge \gamma(w, P_3)$$

in order to define the join $X \bigvee K$ where K is a clique.

Corollary 5.4. $\{(X, K, G) \colon K \approx K_m$ *and* $G \approx X \bigvee K_m\}$ *is definable.*

An induced subgraph $A \leq G$ is called a *maximal connected component* precisely when A is connected and if $A < B \leq G$, then B is disconnected; in particular, a maximal connected component is a connected component. For example, if A and B are connected with $A \leq B$, then $G \approx A + B$ has only B as a maximal connected component.

Proposition 5.5. *The set of connected graphs is definable.*

Proof. The claim is that G is connected if and only if there does not exist $B < G$ such that for all E, $B \prec E \leq G$ implies $E \approx B + N_1$.

Clearly, if G is disconnected with $G \approx B + H$ where B is a maximal connected component, then every cover F of B in G is of the form $F \approx B + N_1$.

If G is connected, then for every $B < G$ there exists $x \in G$ with $x \notin B$ but is adjacent to the connected component of B with largest cardinality. Then the induced subgraph on $B \cup \{x\}$ is certainly not isomorphic to $B + N_1$. ∎

Since the property of being connected is definable, we can recognize the maximal connected components.

Lemma 5.6. $\{(A, G) \colon$ *A is a maximal connected component of G*$\}$ *is definable.*

Proof. From Proposition 5.5 and by the definition of maximal connected component. ∎

The following lemma is the first step in showing the definability of the disjoint sum operation; however, it is such a specialized instance of a sum that we must do a little more preparation before we tackle the general case in Section 7.

Lemma 5.7. $\{(A, B, G) \colon G \approx A + B,$ *A, B connected and incomparable* $\}$ *is definable.*

Proof. The claim is that (A, B, G) is in the relation if and only if

(1) A and B are connected and incomparable,
(2) A and B are maximal connected components of G,

and G is smallest under $\leq$ among graphs satisfying (2). ∎

The following sum will be useful in Section 9.

Lemma 5.8. $\{(C,D,\Gamma)\colon C \approx C_m, D \approx C_n$ *for* $n > m > 5,\ \Gamma \approx \sum_{k=m}^{n} C_k\}$ *is definable.*

Proof. The claim is that $\Gamma \approx \sum_{k=m}^{n} C_k$ if and only if every circuit C such that $C_m \leq_* C \leq_* C_n$ is a maximal connected component of Γ, and Γ is smallest under $\leq$ with this property.

Since distinct circuits are incomparable, the argument from Lemma 5.7 can be applied here to establish the result. ∎

6. Martians and Other Useful Constructions

In this section, we will develop the definability of several constructions which will prove most useful in attaining definability of the disjoint sum operation in Section 7.

Definition 6.1. *For* $n \geq 1$*, a* martian $M(n)$ *is constructed from the two graphs* K_n *and* $K_{1,3}$ *by identifying a single vertex of* K_n *with a single vertex of* $K_{1,3}$ *which has degree one. Note the choice of the vertex in* K_n *is immaterial and so the construction is well defined on isomorphism types. A* p-martian*, denoted by* $pM(n)$*, is constructed from* $M(n)$ *by connecting the remaining two vertices of degree one in* $K_{1,3}$*; thus connecting the "antennae".*

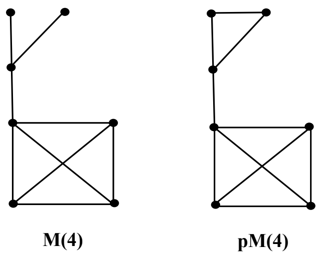

Figure 2: A martian and p-martian.

We start with several preliminary constructions.

Lemma 6.2. $\{(K,\Gamma)\colon K \approx K_n$ *and* $\Gamma \approx K_n \rightarrow_1\}$ *is definable.*

Proof. The claim is that $\Gamma \approx K_n \rightarrow_1$ if and only if

$n = 1$ and $\Gamma \approx K_2$, or
$n > 1$ and Γ covers K_n, Γ is connected, but Γ has a disconnected subcover.

To see this, suppose Γ satisfies the conditions and $n > 1$. Then $K_n \prec \Gamma$ implies $\Gamma \approx K_n +_k N_1$ for some $1 \leq k \leq n$. Since Γ is connected we must have $k \geq 1$. If $k \geq 2$, then every lower cover of Γ is connected which yields a contradiction; therefore, $k = 1$. That the conditions are necessary is immediate. ∎

In the same way we have the definability of the graphs $C_n \to_1$.

Lemma 6.3. *$\{(C,\Gamma)\colon C \approx C_n$ and $\Gamma \approx C_n \to_1\}$ is definable.*

Lemma 6.4. *$\{(C,\Gamma)\colon C$ is a circuit and $\Gamma \approx C + K_2\}$ is definable.*

Proof. The claim is that $\Gamma \approx C_m + K_2$ if and only if

$m = 3$ and $\Gamma \approx K_3 + K_2$, or
$m > 3$ and there exists F such that $C_m \prec F \prec \Gamma$, $\alpha(\Gamma) = \alpha(C_m) + 1$, and C_m is a maximal component of Γ.

It is straightforward to see that $C_m + K_2$ satisfies the criteria.

For sufficiency, suppose Γ satisfies the conditions and $\Gamma \not\approx K_3 + K_2$. We may construct Γ from C_m by adding two new vertices u and v, and possibly some new edges. Since C_m is maximal component of a subcover, u and v are not connected to any vertices of C_m. The condition $\alpha(\Gamma) = \alpha(C_m) + 1$ implies $u \sim v$; therefore, $\Gamma \approx C_m + K_2$. ∎

The graph $C \to_2$ refers to the cover of $C \to_1$ formed by adding an additional vertex and only one additional edge joining the new vertex to the unique vertex of degree one in $C \to_1$. The graph $K_n \to_2$ is defined in a similar manner.

Lemma 6.5. *$\{(C,\Gamma)\colon C$ is a circuit and $\Gamma \approx C \to_2\}$ is definable.*

Proof. The claim is that $\Gamma \approx C \to_2$ where C is a circuit if and only if

(1) $C + N_1 \prec \Gamma$,
(2) Γ is connected,
(3) Γ has a disconnected acyclic subcover.

We shall only verify sufficiency. Suppose Γ satisfies conditions (1)–(3). Then Γ may be constructed from $C + N_1$ by adding an additional vertex v and possibly some new edges joining v to C. If u denotes the isolated vertex of $C + N_1$, then condition (2) implies we have edges $v \sim u$ and $v \sim x$ for some $x \in C$. If v is adjacent to any other vertices of C, then every acyclic subcover is connected, a contradiction of (3); thus, v is adjacent to only one vertex of C which implies $\Gamma \approx C \to_2$. ∎

Proposition 6.6. *$\{(K,\Gamma)\colon K \approx K_n$ and $\Gamma \approx K_n \to_2\}$ is definable.*

Proof. The claim is that $\Gamma \approx K_n \to_2$ if and only if

$n = 1$ and $\Gamma \approx P_3$, or
$n = 2$ and $\Gamma \approx P_4$, or
$n \geq 3$ and $K_n \to_1 \prec \Gamma$, $K_n + N_1 \prec \Gamma$, $P_4 \leq \Gamma$, $K_{n+1} \not\leq \Gamma$, and $C_4 \not\leq \Gamma$.

It is easy to see these conditions are satisfied by $K_n \to_2$.

Suppose Γ meets these conditions, and we may assume $n \geq 3$. Since $K_n + N_1 \prec \Gamma$, we can construct Γ from $K_n + N_1$ by adding a new vertex v and possibly new edges of the form $x \sim v$ where $x \in K_n + N_1$. Since $K_{n+1} \not\leq \Gamma$, there exists $y \in K_n$ such that $y \not\sim v$. Since $K_n \to_1 \prec \Gamma$, there exists $u \in K_n$ such that $u \sim v$, and v is not adjacent to any other vertex of K_n. If z denotes the solitary vertex of N_1, then the condition $P_4 \leq \Gamma$ implies $v \sim z$, and this demonstrates that $\Gamma \approx K_n \to_2$. ∎

Definition 6.7. *Given two connected graphs A and B, the* pointed sum *is a graph* $A +_p B$ *formed by adding a new vertex* v *to* $A + B$ *and two new edges incident to* v*; one edge connects* v *to a vertex in A, and the other edge connects* v *to a vertex in B. Different choices of vertices in A and B lead to non-isomorphic graphs which are still considered as pointed sums; therefore, the notation* $A +_p B$ *will refer to the finite family of pointed sums for the different choices of vertices in A and B which are adjacent to the added vertex which has degree two.*

We can see that when both A and B are complete, or both are circuits, then the choices of vertices in the definition is immaterial, and in these cases the family of pointed sums collapses to a unique graph. While we do not have definability of general pointed sums, we do have definability in certain useful cases.

Lemma 6.8. $\{(C, D, \Gamma)\colon$ *C, D are incomparable circuits and* $\Gamma \approx C +_p D\}$ *is definable.*

Proof. The claim is that $\Gamma \approx C_n +_p C_m$ if and only if

(1) $C_n + C_m \prec \Gamma$,
(2) Γ is connected,
(3) Γ has a disconnected subcover $E \prec \Gamma$ such that $C_n \nleq E$,
(4) Γ has a disconnected subcover $F \prec \Gamma$ such that $C_m \nleq E$.

Suppose Γ satisfies conditions (1)–(4). Then (1) implies Γ may be constructed by adding an additional vertex v to $C_n + C_m$ and some new edges incident with v. Since Γ is connected, there must be edges connecting v to C_n and v to C_m. Take $n \geq m$ and suppose there are at least two vertices in C_n adjacent to v, then no matter how v is connected to C_m, every subcover which avoids C_n is connected; a contradiction to (3). A similar argument shows that v is adjacent to only one vertex of C_m; therefore, $\Gamma \approx C_n +_p C_m$.

For necessity, it is easy to see the only subcovers of $C_n +_p C_m$ which avoid C_n or C_m are $P_{n-1} + C_m$ or $C_n + P_{m-1}$, respectively. ∎

Lemma 6.9. $\{(C, D, G)\colon$ *C, D are incomparable circuits and* $G \approx C \rightarrow_1 + D\}$ *is definable.*

Proof. The definability of this relation follows from Lemmas 5.7 and 6.3. ∎

Lemma 6.10. *The relation*

$$\{(C, D, G)\colon \textit{C and D are incomparable circuits and } G \approx C \rightarrow_1 + D \rightarrow_1\}$$

is definable.

Proof. Notice that when C and D are incomparable circuits, $C \rightarrow_1$ and $D \rightarrow_1$ must also be incomparable. The result now follows from Lemma 5.7. ∎

Let $\gamma(n, m)$ denote the graph formed by adding a single new edge connecting the two unique vertices of $C_n \rightarrow_1 + C_m \rightarrow_1$ which have degree one.

Proposition 6.11. *The relation*

$$\{(C, D, \Gamma) \colon C, D \text{ are incomparable circuits and } \Gamma \approx \gamma(|C|, |D|)\}$$

is definable.

Proof. The claim is that $\Gamma \approx \gamma(n, m)$ if and only if

(1) $C_n \rightarrow_1 + C_m \prec \Gamma$,
(2) Γ is connected,
(3) if $C_m \prec R \leq \Gamma$, then $R \approx C_m \rightarrow_1$ or $R \approx C_m + N_1$,
(4) if $C_n \rightarrow_1 \prec R \leq \Gamma$, then $R \approx C_n \rightarrow_2$, or $R \approx C_n \rightarrow_1 + N_1$.

We only verify sufficiency. Suppose Γ satisfies conditions (1)–(4). From condition (1), we can form G by adding an additional vertex v to $C_n \rightarrow_1 + C_m$ and possibly some new edges incident with v. Since Γ is connected by (2), there is at least one new edge connecting v to the copy of C_m in $C_n \rightarrow_1 + C_m$, but by condition (3) there exists exactly one such edge. Again by (2), there is at least one edge connecting v to $C_n \rightarrow_1$; however, by condition (4), $C_n \rightarrow_2$ is the only possibility for a connected induced subgraph on the vertices of $C_m \rightarrow_1 \cup \{v\}$. We have shown $\Gamma \approx \gamma(n, m)$. ∎

Before we show the definability of martians and p-martians, we need to show the definability of two auxiliary families of graphs.

Lemma 6.12. $\{(K, R, G) \colon K \approx K_n, R \approx K_m, G \approx K_n + K_m\}$ *is definable.*

Proof. We take $n \geq m$. The claim is that $G \approx K_n + K_m$ if and only if G is a disjoint sum of cliques with $\omega(G) = n$, $\alpha(G) = 2$, and

(1) if $n = m$, then G has a unique lower cover, or
(2) if $n > m$, then there exists $B < G$ such that B has a unique lower cover with $\omega(B) = m$ and $\alpha(B) = 2$, and whenever $A < G$ such that $\alpha(A) = 2$ and A has a unique lower cover, then $\omega(A) \leq m$.

If G satisfies the conditions, then $G \approx K_n + K_r$ for some $r \leq n$. Notice G has a unique lower cover exactly when $n = r$. This is the content of condition (1). Notice $K_t + K_t < G$ exactly when $t \leq m$. This is the content of condition (2). ∎

Lemma 6.13. $\{(K, G) \colon K \approx K_n \text{ and } G \approx K_n + P_3\}$ *is definable.*

Proof. The claim is that $G \approx K_n + P_3$ if and only if

$n = 1$ and $G \approx P_3 + N_1$, or
$n = 2$ and $G \approx K_2 + P_3$, or
$n > 2$ and (K_n, P_3, G) is in the relation of Lemma 5.7. ∎

Proposition 6.14. $\{(K, M) \colon K \approx K_n \text{ and } M \approx M(n)\}$ *is definable.*

Proof. The claim is that $M \approx M(n)$ for $n \geq 2$ if and only if

$n = 1$ and $M \approx K_{1,3}$,

$n = 2$ and M is the unique cover of $K_{1,3}$ which is acyclic and embeds P_4, or
$n \geq 3$ and

(1) $K_n + N_2 \prec M$, $K_n \to_2 \prec M$, and these are the only lower covers of M which embed K_n,
(2) $P_4 \leq M$,
(3) $K_{n-1} + P_3 \prec M$.

To show these conditions characterize martians we need only check the cases for $n > 1$.

Suppose M satisfies the conditions and $n = 2$. We may construct M by adding a new vertex v to $K_{1,3}$ and possibly new edges incident with v. Let x be the unique vertex of $K_{1,3}$ with degree three, then $v \sim x$ implies $P_4 \not\leq K_{1,3}$ no matter what other edges are present; thus, $v \not\sim x$. Again, since $P_4 \leq M$, v is adjacent to a vertex of $K_{1,3}$ with degree one, but v must be adjacent to exactly one such vertex since M is acyclic. This finishes the demonstration that $M \approx M(2)$.

Suppose $n \geq 3$ and M satisfies (1)–(3). Since $K_n + N_2 \prec M$, we may construct M by adding an additional vertex v to $K_n + N_2$ and possibly some new edges incident with v. Let w and x be the vertices comprising this copy of N_2. By (2), we must have a new edge $a \sim v$ for some $a \in K_n$ and, without loss of generality, an edge $v \sim x$. If there exists $b \in K_n$ such that $b \neq a$ and $b \sim v$, there is no possibility for $K_{n-1} + P_3 \prec M$, a contradiction to (3); thus, the induced subgraph on $K_n \cup \{v\}$ is isomorphic to $K_n \to_1$. Suppose $w \not\sim v$. Then M has the three lower covers $K_n \to_2$, $K_n + N_2$, and $K_n \to_1 + N_1$ which all embed K_n, a contradiction to (1); therefore, we must have an edge $w \sim v$ and conclude that $M \approx M(n)$.

For necessity, it is easy to see that $M(2)$ is an acyclic cover of $K_{1,3}$ which embeds P_4.

Since $M(n)$ has a unique copy of K_n for $n \geq 3$, the only lower covers of $M(n)$ which embed K_n must come from deleting the vertices not included in K_n. In this case, $K_n \to_2$ and $K_n + N_2$ are the only such covers. Conditions (2) and (3) are immediate. Altogether we have shown (1)–(3) characterize these martians. ∎

Proposition 6.15. *$\{(K, M)\colon K \approx K_n$ and $M \approx pM(n)\}$ is definable.*

Proof. The claim is that $M \approx pM(n)$ if and only if

$n = 1$ and $M \approx K_3 \to_1$, or
$n = 2$ and $K_3 \to_2$, or
$n \geq 3$ and $K_n + K_2 \prec M$, $P_4 \leq M$, $P_5 \not\leq M$, and $\mathbf{D} \not\leq M$.

It is straightforward to check that each of these conditions must hold for the appropriate p-martian, so we shall concentrate on demonstrating that they are sufficient to characterize these graphs. Suppose M satisfies the conditions, and we may assume $n > 2$.

Then $K_n + K_2 \prec M$ implies we may construct M by adding an additional vertex x to $K_n + K_2$ and possibly some new edges. Since $P_4 \leq M$, but $P_5 \not\leq M$, we must have that x is adjacent to each vertex of K_2, and that x is adjacent to at least one vertex $a \in K_n$ and there exists $b \in K_n$ such that $x \not\sim b$. If x is adjacent to an additional vertex $c \in K_n$ distinct from a, then the induced subgraph on vertices $\{c, b, a, x\}$ is isomorphic to **D**, a contradiction; thus, there are no additional edges and we see that $M \approx pM(n)$. This finishes the proposition. ∎

7. G + H

In this section we will prove the definability of the operation $G+H$ where G and H are arbitrary graphs. We start by showing the definability of a special family of pointed sums.

Lemma 7.1. *The relation*

$$\big\{(A, K, G)\colon A \text{ connected and not a clique, } K \approx K_n,\ n > \omega(A)+1,\ G \in A +_p K\big\}$$

is definable.

Proof. Note that A is not complete and $n > \omega(A)+1$ implies that A and K_n are incomparable, and so by Lemma 5.7, $A+K_n$ is definable.

The claim is that (A, K, G) is in the relation if and only if

(1) A is connected and not a clique,
(2) $K \approx K_n$ for some n such that $n > \omega(A)+1$,
(3) $A+K_n \prec G$,
(4) G is connected,
(5) if $K_n \prec R \leq G$, then $R \approx K_n + N_1$ or $R \approx K_n \rightarrow_1$,
(6) $M(n) \nleq G$ and $pM(n) \nleq G$.

The proof of necessity is straightforward and so we will establish sufficiency.

Suppose G satisfies the criteria. By (3), G may be constructed from $A+K_n$ by adding a new vertex p and perhaps some new edges incident with p. Note $n \geq 3$. Let V be the induced subgraph on the vertices of K_n together with p. Since G is connected, the vertex p is connected to some vertex of this copy of K_n; thus, $R \approx K_n \rightarrow_1$. Again, since G is connected there is at least one new edge $p \sim a$ with $a \in A$. Suppose $p \sim b$ where $b \in A$ and $b \neq a$. If $b \sim a$, then the induced subgraph on $K_n \cup \{p, a, b\}$ is isomorphic to $M(n)$, a contradiction. If $b \not\sim a$, then we have a copy of $pM(n) \leq G$, another contradiction; therefore, p is adjacent to at most one vertex of A, and so we conclude that $G \in A +_p K_n$. ∎

Proposition 7.2. $\{(A, \Gamma)\colon$ *A is connected and* $\Gamma \approx A+A\}$ *is definable.*

Proof. If A is a clique, then the definability of $A+A$ is guaranteed by Lemma 6.12. We may suppose A is not a clique, and thus, $|A| > 2$. Let $m = \omega(A)+3$.

Since A and K_m are connected and incomparable, $A+K_m$ is definable by Lemma 5.7, and from Corollary 5.4, $(A+K_m)\bigvee N_1$ is definable. Then by Lemma 7.1, we have that $A +_p K_{m+1}$ is definable where, in a slight abuse of notation, $A +_p K_{m+1}$ will refer to any of the graphs in the set represented by the pointed sum $A +_p K_{m+1}$. By choice of m it is easy to see that $(A+K_m)\bigvee N_1$ and $A +_p K_{m+1}$ must be incomparable; therefore, by Lemma 5.7 the disjoint union $(A+K_m)\bigvee N_1 + (A +_p K_{m+1})$ is definable.

We claim that $A+A+K_{m+1}+K_m$ is the unique graph Γ so that there exists E such that

(1) $\Gamma \prec E \prec (A+K_m)\bigvee N_1 + (A+_p K_{m+1})$,
(2) $K_m \rightarrow_1 \not\leq E$ and K_{m+1} is a maximal connected component of E,
(3) A and K_{m+1} are the only maximal connected components of Γ.

To see this, set $G \approx (A+K_m)\bigvee N_1 + (A+_p K_{m+1})$. By (1), there exist vertices z and w such that $\Gamma \approx G - z - w$. In the construction of $A +_p K_{m+1}$, a new vertex v was added to the sum $A + K_{m+1}$ and an edge connecting v to a vertex of K_{m+1}. Let a denote this vertex of K_{m+1}. If $a \in \{z, w\}$, then K_{m+1} cannot appear as a maximal connected component; a contradiction to (3). If $v \notin \{z, w\}$, then $K_m \rightarrow_1$ embeds in every subcover of G, or K_{m+1} is not a maximal connected component; a contradiction to (2). Without loss of generality, we may take $v = z$. If we let q denote the unique vertex of N_1 in the construction of $(A+K_m)\bigvee N_1$ which is connected to every vertex of $A+K_m$, then (3) implies we must have $q = w$. Then $G - v - q \approx A + A + K_{m+1} + K_m$.

We will now see how to recover $A+A$ from $A+A+K_{m+1}+K_m$. This is the purpose of the following claim which will complete the proposition.

Claim. Consider the following property for a graph H:

$$H + N_2 \leq A + A + K_{m+1} + K_m \quad \text{but} \quad H + N_3 \not\leq A + A + K_{m+1} + K_m. \qquad (**)$$

The graph $A+A$ is the unique graph among those maximal under $\leq$ for property $(**)$, which have A as the only maximal connected component.

Proof. Let X be maximal for property $(**)$ and have A as the only maximal connected component. We may write $X \approx A + X_2 + \cdots + X_n$ where X_i are the connected components of X. Since $X + N_2 \leq A + A + K_{m+1} + K_m$, it must be the case that $X_2 + \cdots + X_n + N_2 \leq A + K_{m+1} + K_m$. Let $G_2 + \cdots + G_n + U$ be an induced subgraph of $A + K_{m+1} + K_m$ such that each $G_i \approx X_i$ and $N_2 \approx U$, and $X_2 + \cdots + X_n + N_2 \approx G_2 + \cdots + G_n + U \subseteq A + K_{m+1} + K_m$. The graph $G_2 + \cdots + G_n + U$ fixes a copy of $X_2 + \cdots + X_n + N_2$ in $A + K_{m+1} + K_m$.

If $U \subseteq K_{m+1} + K_m$, then $G_2 + \cdots + G_n \subseteq A$ which implies $n = 2$ by maximality; therefore, $G_2 \approx A$ and so $X \approx A + A$. We show this is the only possible case.

If $U \subseteq A + K_{m+1}$ but $U \not\subseteq A$, then $G_2 + \cdots + G_n \subseteq A + K_m$. Only a single connected component G_i can be in an induced subgraph in K_m, and so by maximality there is a component, say G_2, isomorphic to K_m. This implies $G_3 + \cdots + G_n \subseteq A$. But $m = \omega(A) + 3$ implies that K_m is a maximal connected component of X; a contradiction. The same argument for $U \subseteq A + K_m$ shows there exists some $G_i \approx K_{m+1}$ which yields another contradiction.

If $N_2 \subseteq A$, then some $G_i \subseteq K_{m+1}$ and maximality again shows we must have some $G_i \approx K_{m+1}$; a contradiction.

This finishes the proof of the claim and the proposition. ∎

Lemma 7.3. $\{(A, B, \Gamma) : A, B$ *connected*, $A < B,\ \Gamma \approx A + B\}$ *is definable.*

Proof. If B is a clique, then so is A and we already have definability of their sum. Assume B is not a clique and set $m = \omega(B) + 3$. Then any two graphs from $A +_p K_{m+1}$ and $B +_p K_m$ are incomparable. By the same argument as in the previous lemma, we have that the sum $\Gamma \approx A + B + K_{m+1} + K_m$ is definable. We can then recover $A + B$ from Γ in a similar way, as well.

$A+B$ is the unique graph H which has B as the only maximal connected component and is maximal under $\leq$ for the property that $H+N_2 \leq \Gamma$ and $H+N_3 \nleq \Gamma$. To see this, let $H \approx H_1 + \cdots + H_r$ be such a graph with connected components H_i. We may take $H_1 \approx B$ and consider $B+H_2+\cdots+H_r+N_2 \leq A+B+K_{m+1}+K_m$. Let $V+G_2+\cdots+G_n+U$ be an induced subgraph of $A+B+K_{m+1}+K_m$ such that each $G_i \approx H_i$, $N_2 \approx U$, and $V \approx B$, and $B+H_2+\cdots+H_n+N_2 \approx V+G_2+\cdots+G_n+U \subseteq A+B+K_{m+1}+K_m$.

If $N_2 \approx U \nsubseteq K_{m+1}+K_m$, then by maximality some component G_i must intersect $K_{m+1}+K_m$. Again by maximality, we can conclude that $H_i \approx K_m$ or $H_i \approx K_{m+1}$; a contradiction. It must be that $U \subseteq K_{m+1}+K_m$ and intersects both cliques which implies $G_2+\cdots+G_r \leq A$; thus, maximality implies $r=2$ and so $H_2 \approx A$. Altogether, this shows $H \approx A+B$. ∎

Using the last two results and Lemma 5.7, we have definability of the following relation which captures when a graph is a disjoint sum of two connected graphs.

Proposition 7.4. $\{(A,B,\Gamma) : A, B$ *are connected and* $\Gamma \approx A+B\}$ *is definable.*

We should note a useful property of the join construction. If V is a disconnected graph, then $V \bigvee N_1$ is connected and has a unique disconnected subcover; namely, if $U \prec V \bigvee N_1$ is disconnected, then $U \approx V$.

Lemma 7.5. *The relation*

$$\{(U,V,\Gamma) : V \text{ disconnected}, U \text{ a maximal connected component of } \Gamma, \Gamma \approx U+V\}$$

is definable.

Proof. (U,V,Γ) is in the relation if and only if

(1) V is disconnected and $V \leq \Gamma$,
(2) U is a maximal connected component of Γ,
(3) if M is a maximal connected component of Γ, then $M \approx U$ or $M \leq V$,
(4) $\Gamma \prec U+V \bigvee N_1$,
(5) Γ is not isomorphic to the disjoint sum of two connected graphs.

Since necessity is straightforward to check, we only prove sufficiency.

Suppose Γ satisfies conditions (1)–(5). By condition (4), $\Gamma \approx \Gamma'$ where Γ' is an induced subgraph of $U+V \bigvee \{v\}$, and there exists a vertex $x \in U+V \bigvee \{v\}$ such that $U+V \bigvee \{v\} - x = \Gamma'$. We will show Γ' is isomorphic to $U+V$ by considering the possible choices for the vertex x.

Suppose $x=v$. Then $V \bigvee \{v\} - x = V$ and so $\Gamma' = U+V \bigvee \{v\} - x = U+V$.

If $x \in V$, then $V \bigvee \{v\} - x$ is connected. This implies $\Gamma' = U+V \bigvee \{v\} - x$ is the disjoint sum of two connected graphs, a contradiction to condition (5). Therefore, $x \notin V$.

Suppose $x \in U$. Since $V \bigvee \{v\}$ is a connected subset of Γ', we must have that $V \bigvee \{v\} \leq M$ for some maximal connected component of $\Gamma' \approx \Gamma$. By condition (3), we have $V \bigvee \{v\} \leq M \approx U$ or $V \bigvee \{v\} \leq M \leq V$. Since $|V \bigvee \{v\}| > |V|$, we must have $V \bigvee \{v\} \leq U$. But $\Gamma' = (U-x) + V \bigvee \{v\}$ and condition (5) implies $U-x$ is disconnected. Since $|U-x|<|U|$, condition (2) implies $U \leq V \bigvee \{v\}$ which then yields $U \approx V \bigvee \{v\}$.

Since $U-x$ is disconnected, then $U-x \approx V \bigvee \{v\} - x \approx V$ implies $\Gamma' = (U-x)+(V \bigvee \{v\}) \approx V+U$. ∎

Lemma 7.6. $\{(A,P,\Gamma): P \text{ a path, } P \nleq A, \text{ and } \Gamma \approx A+P\}$ *is definable.*

Proof. If A is a path or is just connected, then we already have the definability of $A+P$. If A is disconnected, then the definability of $A+P$ follows from Lemma 7.5 since the condition $P \nleq A$ implies that P is a maximal connected component of $A+P$. ∎

Proposition 7.7. $\{(A,B,\Gamma): \Gamma \approx A+B\}$ *is definable.*

Proof. Let P be a path such that $P \nleq A$ and $P \nleq B$ and $|P|>3$. Set $H=(A+P)\bigvee N_1 + (B+P)\bigvee N_1$ which is definable from A and B using Lemma 7.6, Proposition 7.4, and Corollary 5.4. The claim is that $A+B+P+P$ is the unique graph G such that

(1) $G \prec E \prec H$ for some E,
(2) P is a maximal connected component of G,
(3) $P+P \leq G$,
(4) $P \bigvee N_1 \nleq G$.

To see this, assume G satisfies conditions (1)–(4). We can write $H = A' \cup B' \cup P' \cup P'' \cup \{p,q\}$ where $A' \approx A$, $B' \approx B$, $P' \approx P'' \approx P$, and $A' \cup P' \cup \{p\} \approx (A'+P')\bigvee N_1$ and $B' \cup P'' \cup \{q\} \approx (B'+P'')\bigvee N_1$. Then by (1), $G = H - \{u,v\}$ for some vertices u, v. If neither p nor q is in $\{u,v\}$, then P cannot be a maximal connected component of G; a contradiction to condition (2). We may assume, without loss of generality, $q=v$. Then $G=(A'+P')\bigvee N_1 + B' + P'' - u$. If $u \in B'$ or $u \in P''$ or $u \in A'$, then we have $P \bigvee N_1 \nleq G$ which contradicts condition (4). If $u \in P'$, then $P+P \nleq G$ since any copy of P in $(A'+P')\bigvee N_1 - u$ will contain p and two consecutive vertices of P' which induce a copy of K_3; a contradiction of condition (3). It must be the case that $u=p$, which implies

$$G=(A'+P')\bigvee N_1 + B' + P'' - p = A'+P'+B'+P'' \approx A+B+P+P.$$

We may then use Lemma 7.6 to capture $A+B$ as the unique graph F such that $F+P+P \approx G$. ∎

8. Cardinality

In this section, we establish that the cardinality of graphs is a definable property.

Proposition 8.1. $\{(K,A): K \approx K_n \text{ and } |A|=n\}$ *is definable.*

Proof. It suffices to characterize when $|A| \geq n$. The claim is that $|A| \geq n$ if and only if A is a clique and $K_n \leq A$, or A is not a clique, $\omega(A)=m$, and for every graph P with the following properties we must have $K_{n+1} \leq P$:

(1) $A \bigvee K_1 \leq P$,

(2) for every clique K and graph Q such that $K_m < Q \leq A$ and $Q \bigvee K \leq P$, there exist Q' such that $K_m \leq Q' \prec Q$ and a clique $\overline{K}$ with $K \prec \overline{K}$ such that $Q' \bigvee \overline{K} \leq P$.

Assume first that A is not a clique and that $|A| \geq n$. Assume P satisfies conditions (1)–(3). By induction on k for $0 \leq k \leq |A| - m$, we argue that $Q \vee K_{k+1} \leq P$ for some graph Q such that $K_m \leq Q \leq A$ with $k = |A| - |Q|$. We see that (1) yields the base case $k = 0$ with $Q = A$. Condition (3) is applied at the inductive step for $1 \leq k < |A| - m$ to show $Q' \vee K_{k+1} \leq P$ for some $K_m \leq Q'$. At the step $k = |A| - m = |A| - |Q|$, we see that $K_m \leq Q$ implies $K_m \approx Q$ which yields

$$P \geq Q \vee K_{k+1} \approx K_m \vee K_{|A|-m+1} \approx K_{|A|+1} \geq K_{n+1}.$$

Now assume A is not a clique and $|A| < n$. Let $Q_1, \ldots, Q_p$ be a full list, up to isomorphism, of all graphs Q such that $K_m \leq Q \leq A$. Let $r_i = |A| - |Q_i|$. We see that $\omega(Q_i) = m$. Set

$$P = \sum_{i=1}^{p} Q_i \vee K_{r_i+1}.$$

Note $\omega(Q_i \vee K_{r_i+1}) = m + r_i + 1$ and $\omega(P) = \max\{\omega(Q_i \vee K_{r_i+1})\colon i \leq p\}$. The clique size of P is determined by the component of maximum clique size which occurs when r_i is largest; that is, when $|Q_i| = |K_m| \Rightarrow Q_i \approx K_m$. For simplicity, let this occur at $i = 1$ and so we have

$$Q_1 \vee K_{r_1+1} \approx K_m \vee K_{|A|-m+1} \approx K_{|A|+1}.$$

Thus, $\omega(P) = |A| + 1 \leq n \Rightarrow K_{n+1} \nleq P$. It remains to show that P satisfies (1) and (2).

Since A is not a clique, condition (1) is immediately seen to hold by construction. For (2), suppose $K_m < Q \leq A$ and K is a clique such that $Q \vee K \leq P$. Since $Q \vee K$ is connected, we must have $Q \vee K \leq Q_i \vee K_{r_i+1}$ for some $i \in [p]$. Let $K \approx K_l$. Then

$$\omega(Q) + l = \omega(Q \vee K) \leq \omega\left(Q_i \vee K_{r_i+1}\right) = m + r_i + 1$$

implies $l \leq r_i + 1$. We can assume $Q \vee K \subseteq Q_i \vee K_{r_i+1}$.

Since $K_m < Q \leq A$, Q is not a clique, and so there exist $U \subseteq Q$ and $q_0, q_1 \in Q$ such that $K_m \approx U$ and $q_0 \not\sim q_1$. We may take, without loss of generality, $q_0 \notin U$. We must have $\{q_0, q_1\} \subseteq Q_i - K_{r_i+1}$. So there exists $j \neq i$ such that $Q_j \approx Q_i - q_0$. Put $Q' = Q - q_0$. Then $K_m \leq Q'$ and $Q' \vee K_l \subseteq (Q_i - q_0) \vee K_{r_i+1} \approx Q_j \vee K_{r_i+1}$. Clearly, we have $K_{r_j+1} \approx K_{r_i+1} \vee K_1$. Set $K_l \prec K_l \vee K_1 = \overline{K}$. To finish the proposition we see that

$$Q' \vee \overline{K} \approx Q' \vee (K_l \vee K_1) \leq Q_j \vee (K_{r_i+1} \vee K_1) \approx Q_j \vee K_{r_j+1}. \qquad \blacksquare$$

We can use Proposition 8.1 to define the pairs (K, N) of cliques and trivial graphs with the same cardinality. Since addition with trivial graphs is accomplished by Corollary 4.5, addition with the cardinality of arbitrary graphs is given by the ternary relation $\{(A, B, G)\colon |A| + |B| = |G|\}$. As a consequence, we have the definability of the n-step cover $\prec_n$ defined as $A \prec_n B$ if there exists a chain of covers $A \prec F_1 \prec \cdots \prec F_n \approx B$.

Lemma 8.2. *$\{(A, B, C)\colon C \approx C_n$ and $A \prec_n B\}$ is definable.*

Proof. The claim is that $A \prec_n B$ precisely when $A \leq B$, and $C_n \approx C_{|B|-|A|}$. $\blacksquare$

9. Individual Definability

Most of this section will be taken up with a proof of the following proposition.

Proposition 9.1. *Every graph is definable* (*up to isomorphism*) *in* $\langle \mathcal{G}, \leq, P_3 \rangle$.

As a corollary, we can now establish one of the principal results of this paper.

Theorem 9.2. *Let* $\langle \mathcal{PG}, \leq \rangle$ *denote the ordered set of finite isomorphism types of simple graphs ordered by substructure, and* $\mathcal{P}_3$ *the isomorphism type of the path on 3 vertices. Then every element of* $\langle \mathcal{PG}, \leq, \mathcal{P}_3 \rangle$ *is first-order definable; consequently, edge-complementation* **co** *is the only non-trivial automorphism of* $\langle \mathcal{PG}, \leq \rangle$*, and for each finite isomorphism type G, the set* $\{G, G^{\mathbf{co}}\}$ *is definable without constants.*

Proof. That every finite isomorphism type is definable is precisely the content of Proposition 9.1 after factoring the equivalence relation determined by isomorphism on the quasi-ordered set $\langle \mathcal{G}, \leq, P_3 \rangle$.

We now show **co** is the unique non-trivial automorphism of $\langle \mathcal{PG}, \leq \rangle$. Suppose τ is an automorphism which fixes $\mathcal{P}_3$. By assumption, for every element $G \in \mathcal{PG}$, there is a formula $\Psi_G(x, y)$ such that $\langle \mathcal{PG}, \leq \rangle \models \Psi_G(H, \mathcal{P}_3)$ if and only if $G = H$. Then $\langle \mathcal{PG}, \leq \rangle \models \Psi_G(G, \mathcal{P}_3)$ implies $\langle \mathcal{PG}, \leq \rangle \models \Psi_G(\tau(G), \tau(\mathcal{P}_3))$, which implies $\langle \mathcal{PG}, \leq \rangle \models \Psi_G(\tau(G), \mathcal{P}_3)$; therefore, $\tau(G) = G$ for every $G \in \mathcal{PG}$, and so we conclude $\tau = \mathrm{id}$.

If $\tau \neq \mathrm{id}$, then because $\{\mathcal{P}_3, \mathcal{K}_2 + \mathcal{N}_1\}$ is definable without constants, we must have $\tau(\mathcal{P}_3) = \mathcal{K}_2 + \mathcal{N}_1 = \mathbf{co}(\mathcal{P}_3)$. Then $\mathbf{co}^{-1} \circ \tau$ fixes $\mathcal{P}_3$, and so by the above argument, $\mathbf{co}^{-1} \circ \tau = \mathrm{id}$, which implies $\tau = \mathbf{co}$.

For the last claim in the theorem, let $\Theta(x)$ be the formula which defines $\{\mathcal{P}_3, \mathcal{K}_2 + \mathcal{N}_1\}$ without constants. Then for each $G \in \mathcal{PG}$, $\exists y \Psi_G(x, y) \wedge \Theta(y)$ defines $\{G, G^{\mathbf{co}}\}$ without constants. ∎

The previous theorem has the interesting consequence that every finite graph isomorphic to its edge-complement is definable in the substructure relation without constants. We now turn toward proving Proposition 9.1.

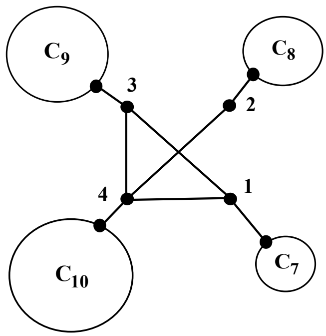

Figure 3: An o-presentation for $P_4(K_3 \rightarrow_1, B)$.

Definition 9.3. *Let $A, B \in \mathcal{G}$ with $B \approx A$ and $|A| = n$. Construct a finite graph denoted by $P_n(A, B)$ in the following way:*

First, take the graph $B + \sum_{i=1}^{n} C_{n+2+i}$. Next, for each vertex k of B, add an edge connecting k to some vertex of C_{n+2+k}. In the end only n new edges are added. The resulting graph is called an o-presentation *of A. The o-presentation $P_n(A, B)$ should look like the graph A with an edge leading out of each vertex to a circuit uniquely determined by cardinality. Figure 3 shows $P_4(A, B)$ where A is the isomorphism type of $K_3 \rightarrow_1$ and B is the isomorphic copy over the positive integers labeled as shown.*

Proposition 9.4. *For a fixed $A \in \mathcal{G}$, each o-presentation $P_n(A, B)$ is definable.*

Proof. The idea is to use specific information of B as a graph on the vertices $[n]$ to write down first-order properties which capture $P_n(A, B)$. First we introduce a little simplifying notation: for $i, j \in [n]$, $i \neq j$, let

$$B(i, j) = \begin{cases} C_{n+2+i} \rightarrow_1 + C_{n+2+j} \rightarrow_1, & \text{if } i \not\sim j \text{ in } B, \\ \gamma(n+2+i, n+2+j), & \text{if } i \sim j \text{ in } B. \end{cases}$$

The claim is that $\Gamma \approx P_n(A, B)$ if and only if

(1) $\sum_{i=1}^{n} C_{n+2+i} \prec_n \Gamma$,
(2) if $C_{n+2+i} \prec R \leq \Gamma$, then $R \approx C_{n+2+i} + N_1$ or $R \approx C_{n+2+i} \rightarrow_1$,
(3) for each $i, j \in [n]$, $i \neq j$, $B(i, j) \leq \Gamma$,
(4) if $\sum_{i=1}^{n} C_{n+2+i} \prec R \leq \Gamma$, then there exists $j \in [n]$ such that $C_{n+2+j} \rightarrow_1 \leq R$,
(5) for each $i, j \in [n]$, $i \neq j$, if $C_{n+2+i} + C_{n+2+j} \prec R \leq \Gamma$, then $C_{n+2+i} +_p C_{n+2+j} \not\approx R$,
(6) for $i \in [n]$, if $C_{n+2+i} \prec_2 R \leq \Gamma$, then $R \approx C_{n+2+i} \rightarrow_2$, or $R \approx C_{n+2+i} + K_2$, or $R \approx C_{n+2+i} \rightarrow_1 + N_1$, or $R \approx C_{n+2+i} + N_2$.

It is easy to see that because the cardinality of C_{n+2+i} exceeds n and is connected uniquely to vertex i of B, $P_n(A, B)$ contains a unique copy of each C_{n+2+i} for $i \in [n]$; therefore, $P_n(A, B)$ also contains a unique copy of $\sum_{i=1}^{n} C_{n+2+i}$. By construction, each C_{n+2+i} is connected to a unique vertex. These facts make it straightforward to check that $P_n(A, B)$ satisfies the stated conditions.

For sufficiency, assume Γ satisfies conditions (1)–(6). From (1), we can assume, after passing to isomorphic induced subgraphs, that there exist vertices $\{v_1, \ldots, v_n\}$ of Γ such that $\sum_{i=1}^{n} C_{n+2+i} = \Gamma - v_1 - \cdots - v_n$. Suppose there exist $k, j \in [n]$ such that v_k is adjacent to more than one vertex of C_{n+2+j}. Then the induced subgraph on the vertices $C_{n+2+j} \cup \{v_k\}$ is not isomorphic to $C_{n+2+j} + N_1$ nor to $C_{n+2+j} \rightarrow_1$; a contradiction to (2). Each vertex v_k is adjacent to at most one vertex of any C_{n+2+i} for $i = 1, \ldots, n$.

Suppose there exist $\{k, i, j\} \subseteq [n]$, $i \neq j$, such that the induced graph on the vertices of $C_{n+2+j} \cup \{v_k\}$ is isomorphic to $C_{n+2+j} \rightarrow_1$, and the induced subgraph on the vertices $C_{n+2+i} \cup \{v_k\}$ is isomorphic to $C_{n+2+i} \rightarrow_1$. Then the induced subgraph on $C_{n+2+i} \cup C_{n+2+j} \cup \{v_k\}$ is isomorphic to $C_{n+2+i} +_p C_{n+2+j}$; a contradiction to (5). We see that if v_k is adjacent to some C_{n+2+j}, then it cannot be adjacent to any other circuit of $\sum_{i=1}^{n} C_{n+2+i}$.

If for each $k \in [n]$, we consider the induced subgraph on the vertices of $\sum_{i=1}^{n} C_{n+2+i}$ $\cup \{v_k\}$, then condition (4) implies v_k is adjacent to some C_{n+2+j}. Altogether we have shown there is a function $\phi\colon [n] \to [n]$ such that for each $k \in [n]$, $C_{n+2+\phi(k)}$ is the unique circuit of $\sum_{i=1}^{n} C_{n+2+i}$ adjacent to v_k. Moreover, the induced subgraph on the vertices yields a copy of $C_{n+2+\phi(k)} \to_1$.

We show ϕ is bijective. Suppose $\phi(i) = \phi(j) = k$. Then the induced subgraph on the vertices of $C_{n+2+k} \cup \{v_i, v_j\}$ is a graph which cannot be isomorphic to any of the four types of graphs listed in condition (6), a contradiction; therefore, ϕ is injective and so a bijection.

Condition (1) then implies there is a unique copy of each C_{n+2+i}. Since each $C_{n+2+\phi(k)}$ is uniquely connected to a single v_k, it is easy to see that $B(\phi(i), \phi(j)) \leq \Gamma$ if and only if $B(\phi(i), \phi(j))$ is isomorphic to the induced subgraph on the vertices of $C_{n+2+\phi(i)} \cup C_{n+2+\phi(j)} \cup \{v_i, v_k\}$. This implies $v_i \sim v_j$ if and only if $B(\phi(i), \phi(j)) \leq \Gamma$. Condition (3) then implies $v_i \sim v_j$ if and only if $\phi(i) \sim \phi(j)$ in B.

If F is the graph Γ induces on the vertices $\{v_1, \ldots, v_n\}$, then what we have shown is that the map $v_i \mapsto \phi(i)$ for $i = 1, \ldots, n$ yields an isomorphism $F \approx B$. Since $C_{n+2+\phi(i)}$ is uniquely connected to v_i by a single edge, the isomorphism ϕ extends to an isomorphism $\Gamma \approx P_n(A, B)$ in the natural way. ∎

The next task is to find a way to "read off" the copy of A sitting inside an o-presentation $P_n(A, B)$. The first step is to return to the topic of paths and isolate particular covers. When attaching a new vertex v to a path P, the choice of $u \in P$ for $u \sim v$ makes a difference. We will use the notation $P_n \to_1^t$ to denote the covers of P_n which are formed by adding a single new edge $u \sim v$ where v is a new vertex and $u \in P_n$ such that u has degree two. Different choices of u lead to non-isomorphic graphs, so the notation $P_n \to_1^t$ refers to the finite family of such graphs for a fixed n.

Lemma 9.5. *$\{(P, \Gamma)\colon P \approx P_n$ and $\Gamma \in P_n \to_1^t\}$ is definable.*

Proof. The claim is that $\Gamma \in P_n \to_1^t$ if and only if $P_n \prec \Gamma$, $K_{1,3} \leq \Gamma$, $P_{n+1} \not\leq \Gamma$, and Γ is acyclic.

If Γ satisfies the conditions, then Γ is formed from P_n by adding a new vertex v and possibly new edges of the form $u \sim v$ for $u \in P_n$. If at least two new edges are added, then a circuit must be formed, and so there is at most one new edge $x \sim v$ with $x \in P_n$. Since $K_{1,3} \leq \Gamma$, there is exactly one new edge. Since $P_{n+1} \not\leq \Gamma$, the degree of x cannot be one. This establishes the lemma. ∎

Lemma 9.6. *$\{(C, \Gamma)\colon P \approx P_n$ and $\Gamma \approx \sum_{i=1}^{n} P_{n+1+i}\}$ is definable.*

Proof. The claim is that $\Gamma \approx \sum_{i=1}^{n} P_{n+1+i}$ exactly when $\Gamma \prec_n \sum_{i=1}^{n} C_{n+2+i}$ and Γ is acyclic. ∎

We now have all the ingredients to finish the proof of Proposition 9.1. From an o-presentation $P_n(A, B)$ we see that $A + \sum_{i=1}^{n} P_{n+1+i}$ is the unique graph G such that

(A) $G \prec_n P_n(A, B)$,
(B) for all $k \in [n]$, $C_{n+2+k} \not\leq G$,
(C) for all $k \in [n]$, G embeds no element of $P_{n+1+k} \to_1^t$,
(D) for all $k \in [n]$, each P_{n+1+k} is a connected component of G,

This follows since conditions (A) and (B) imply G is obtained precisely by deleting exactly one vertex from each C_{n+2+i} for $i \in [n]$. Conditions (C) and (D) imply that each of those vertices must have degree three. We can then recover A as the unique graph H such that $G \approx H + \sum_{i=1}^{n} P_{n+1+i}$. ∎

Remark 9.7. Using a particular $P_3 \in \mathcal{G}$ as a constant we have shown every finite graph is definable up to isomorphism in $\langle \mathcal{G}, \leq, P_3 \rangle$. The same result could be achieved, but perhaps with greater difficulty, if we had chosen another graph C as the constant, provided C is not self-complementary. To see this, notice from the proof of Proposition 2.2, that there is a formula $\beta(x)$ in the language of $\leq$ such that $\langle \mathcal{G}, \leq \rangle \models \beta(E)$ if and only if $E \approx P_3$ or $E \approx K_2 + N_1$. By what we have shown, for any $G \in \mathcal{G}$ there is a formula $\phi_G(x, y)$ in the language of $\leq$ such that $\langle \mathcal{G}, \leq \rangle \models \phi_G(E, P_3)$ if and only if $E \approx G$. It is not hard to see that the unary formula

$$(\exists y)\phi_G(x, y) \wedge \phi_C(C, y) \wedge \beta(y)$$

uniquely defines G.

In the next proposition, we shall see how to capture the pair (A, P) where P is isomorphic to some o-presentation of A. The proof of Proposition 9.4 relied on the fact that we had a fixed graph on hand, and so we could "encode" the edge relation of this fixed graph with a certain packet of formulas. This means different graphs require different packages of formulas to define the edge relations in the o-presentations. Since we have definable access to the cardinality of a graph, and can do arithmetic with circuits, we shall be able to describe a uniform packet of formulas which "encode" the edge relations satisfied by an o-presentation.

Proposition 9.8. *We have the following*:

(1) $\{(A, P_n(A, B))\colon$ *for some* $B \approx A$ *with* $n = |A|\}$ *is definable.*
(2) *If* B *is a graph over the vertices* $[n]$ *with* $B \approx A$ *and* B' *is a graph over the vertices* $[m]$ *with* $B' \approx A'$, *then* $P_n(A, B) \approx P_m(A', B')$ *if and only if* $n = m$ *and* $B = B'$.

Proof. For part (1), the definable conditions are (1)–(6) in Proposition 9.4 together with (A)–(D) from before Remark 9.7, except (3) is substituted with

(3') for each $i, j \in [n]$, $i \neq j$, if $C_{n+2+i} \rightarrow_1 + C_{n+2+j} \rightarrow_1 \not\leq P$, then $\gamma(n+2+i, n+2+j) \leq P$.

Here, the previous fixed n is determined by the arbitrary $|A| = n$.

Any P which satisfies conditions (1)–(6) must be isomorphic to an o-presentation $P_n(E, F)$ for some F. Conditions (A)–(D) then allow us to see that $E \approx A$. The details are left to the reader.

We establish part (2). Clearly, $n = m$ and $B = B'$ imply $P_n(A, B) \approx P_m(A', B')$. Suppose $P_n(A, B) \approx P_m(A', B')$. By using the definition of an o-presentation, $P_n(A, B)$ has $n + \sum_{i=1}^{n}(n+2+i)$ vertices. Since $P_n(A, B)$ and $P_m(A', B')$ have the same cardinality, we must have

$$n^2 + 3n + \frac{n(n+1)}{2} = m^2 + 3m + \frac{m(m+1)}{2},$$

which implies $n = m$, and so B and B' have the same vertices. Because $P_n(A,B)$ and $P_m(A',B')$ must then have a unique copy of each C_{n+2+i} and therefore, a unique copy of each $C_{n+2+i} \rightarrow_1$ for $i \in [n]$, the isomorphism of o-presentations restricts to the identity on B and B'. ∎

10. Positive Definability

For a fixed signature, a first-order formula $\phi(x_1,\dots,x_n)$ is said to be open if it contains no quantifiers. A formula is in *prenex* form if it looks like

$$Q_1 y_1 \cdots Q_m y_m \psi(x_1,\dots,x_n),$$

where each Q_i is a quantifier, some of the y_i's may refer to the variables x_j, and $\psi(x_1,\dots,x_n)$ is a formula in which no quantifiers appear. A standard result guarantees that every formula is logically equivalent to some formula in prenex form, which provides a canonical description for choosing interesting species of formulas. For example, we may define one such species by saying that a formula is *universal* if it is logically equivalent to a prenex formula with only universal quantifiers. Once we have a specified type of sentences, say type T, then we can consider classes of structures which are models of a set of type T sentences; for example, a *universal class* is a class of structures which are the models of a set of universal sentences.

In the paper [8], Ježek and McKenzie introduce the following general situation. Let K be a fixed T-class of structures over a finite signature where T denotes a fixed type of axioms such as equations, quasi-equations, or universal sentences. Let $\mathcal{L}_K$ denote the collection of T-subclasses of K which usually forms a complete lattice ordered by inclusion. We may investigate if any of the following conditions are met by the lattice $\mathcal{L}_K$:

- the finitely generated T-subclasses are definable in the language of lattices, and each such T-subclass is definable up to the automorphisms of $\mathcal{L}_K$;
- the finitely axiomatizable T-subclasses are definable in the language of lattices, and each such T-subclass is definable up to the automorphisms of $\mathcal{L}_K$;
- the classes axiomatizable by a single T-axiom is a definable subset of $\mathcal{L}_K$;
- the only automorphisms of $\mathcal{L}_K$ are the "obvious" ones.

If all of the above conditions are met, we say that the T-theories of K have *positive definability*. For earlier work on similar results, see [4–7] and [15] on definability among general equational theories, and [12] for the open questions concerning semigroups.

For an example of an "obvious" automorphism, let P be a poset. Reversing the direction of the ordering produces a new partial order over the same universe, denoted by $P^{\mathbf{op}}$, where

$$a <_P b \quad \text{iff} \quad a >_{P^{\mathbf{op}}} b.$$

The map $\mathbf{op}\colon P \to P^{\mathbf{op}}$ can be seen to preserve the relation of embedding among posets. This means **op** induces a non-trivial automorphism of the poset of finite isomorphism types ordered by substructure. Moreover, it is a non-trivial automorphism when restricted to lattices and distributive lattices. One of the principal results gathered from the series of papers on substructure ordering is the following:

Theorem 10.1. ([8–11]) *Let $\mathcal{U}$ denote either the class of posets, lattices, or distributive lattices. Let $\langle \mathcal{PU}, \leq \rangle$ denote the poset of finite isomorphism types in the class $\mathcal{U}$ ordered by substructure. There exists a single type $\mathbf{c} \in \mathcal{PU}$, such that every element of $\langle \mathcal{PU}, \leq, \mathbf{c} \rangle$ is first-order definable; moreover,* ***op*** *is the only non-trivial automorphism of $\langle \mathcal{PU}, \leq \rangle$.*

If Sem denotes the poset of finite isomorphism types of meet-semilattices ordered by substructure, then every type is first-order definable in the order relation of Sem; in particular, Sem has no non-trivial automorphisms.

For a class $\mathcal{K}$, $\mathbf{S}(\mathcal{K})$ will denote the class of structures isomorphic to substructures of structures in $\mathcal{K}$. The class $\mathbf{P}_U(\mathcal{K})$ will consist of those structures isomorphic to ultraproducts of structures from $\mathcal{K}$. Universal classes can be characterized in the following manner:

Theorem 10.2. ([1, Thm. 2.20]) *For any class of structures over a fixed signature, the following are equivalent*:

(1) *$\mathcal{K}$ is a universal class.*
(2) *$\mathcal{K}$ is closed under $\boldsymbol{S}$ and $\boldsymbol{P}_U$.*
(3) *$\mathcal{K} = \boldsymbol{SP}_U(\mathcal{K}^*)$ for some class $\mathcal{K}^*$.*

For a universal class $\mathcal{U}$, the universal classes contained in $\mathcal{U}$ may be ordered by containment; moreover, the order is a complete lattice order with meet given by intersection and the join of subclasses $\mathcal{K}$ and $\mathcal{V}$ given as $\mathcal{K} \vee \mathcal{V} = SP_U(\mathcal{K} \cup \mathcal{V})$ by Theorem 10.2.

For posets, meet-semilattices and distributive lattices the universal subclasses are determined by their finite members, and so in a natural way the distributive lattice of order ideals of the substructure poset is isomorphic to the lattice of universal subclasses; under the isomorphism, a particular type in the substructure poset corresponds to the universal class generated by its principal order ideal. Since principal order ideals are first-order definable in the distributive lattice of order ideals, definable relations in the substructure ordering yield definable relations in the lattice of universal subclasses. Positive definability can then be established by Theorem 10.1 and the following characterizations which admit first-order definitions in the language of lattice theory:

- A universal subclass is finitely generated if and only if it is contained in a universal subclass generated by a single structure.
- A universal subclass is finitely axiomatizable if and only if up to isomorphism there are only finitely many finite structures minimal under the substructure ordering among those structures omitted from the subclass.

It should be noted that the underlying order of the structures is immaterial. With no change in the arguments [9, Thm. 8.1], these results can be extended to a general finitely axiomatizable universal class $\mathcal{U}$ which is of finite signature, locally finite, for each natural number N there are only finitely many N-generated structures up to isomorphism, and any finite set of finite structures in $\mathcal{U}$ can be embedded into some other finite structure of $\mathcal{U}$. These conditions are often met in pure relational signatures such as digraphs, and in particular, simple graphs. This yields a broad framework for attaining positive definability provided an analogue of Theorem 10.1 can be established; therefore, our Proposition 9.1 yields positive definability for the universal theories of graphs where complementation affords the unique non-trivial automorphism.

Theorem 10.3. *The universal theories of simple graphs has positive definability. The map induced by complementation is the unique non-trivial automorphism of the lattice of universal subclasses.*

The following answers in this setting a very general problem asked in [8]: for which locally finite universal classes $\mathcal{K}$ is the set of non-finitely generated universal subclasses the union of finitely many principal filters in $\mathcal{L}_{\mathcal{K}}$? For a class of graphs H, $\mathcal{U}(H)$ is the smallest universal class of graphs containing H.

Proposition 10.4. *The set of non-finitely generated universal subclasses of simple graphs is equal to the union of the principal filters generated by the classes* $\mathcal{U}(\{K_m \colon m < \omega\})$ *and* $\mathcal{U}(\{N_m \colon m < \omega\})$.

Proof. Since graphs are locally finite, and universal classes are generated under the closure of ultraproducts and substructure [1, V. 2], a subclass is non-finitely generated if and only if it contains infinitely many non-isomorphic finite graphs, and so finite graphs of arbitrarily large order. By Ramsey's Theorem, a non-finitely generated class of graphs must contain complete or trivial graphs of arbitrarily large finite order. Clearly, $\mathcal{U}(\{K_m \colon m < \omega\})$ and $\mathcal{U}(\{N_m \colon m < \omega\})$ are themselves not finitely generated. ∎

The use of Ramsey's theorem is no accident. Let $\mathcal{K}_{fin}$ denote the class of finite structures in $\mathcal{K}$. It is not to difficult to see that for a class $\mathcal{K}$ of structures of finite signature, we have $\mathcal{U}(\mathcal{K})_{fin} \subseteq S(\mathcal{K})_{fin}$. With this, one can show that Proposition 10.4 is actually equivalent to the classic two-color Ramsey theorem.

11. Maps

We now turn to the task of encoding set functions which will be accomplished with Proposition 11.5. This result will be needed in the next section, where we shall be interested in graph homomorphisms. We start with some auxiliary constructions.

Definition 11.1. *A* panda *is the graph $P(n)$ constructed from $C_n \rightarrow_1$ by adding two additional vertices x and y and only two new edges $x \sim u$ and $y \sim u$ where u is the unique vertex of $C_n \rightarrow_1$ with degree one. A* p-panda, *denoted by $pP(n)$, is formed from $P(n)$ by completing the triangle formed by the panda's arms; that is, by adding the edge $x \sim y$ to $P(n)$.*

Lemma 11.2. $\{(C,F) \colon C \approx C_n$ *and* $F \approx P(n)\}$ *is definable.*

Proof. The claim is that $F \approx P(n)$ if and only if

$n = 3$ and $F \approx M(3)$, or

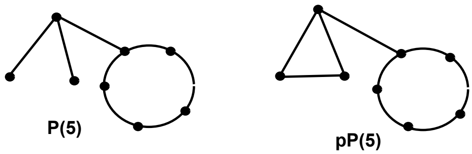

Figure 4: A panda and p-panda.

$n > 3$ and $C_n + N_2 \prec F$, F is connected, and if $C_n \prec R \leq F$, then $R \approx C_n \rightarrow_1$ or $R \approx C_n + N_1$.

Suppose F satisfies the criteria and $F \not\approx M(3)$. We may construct F from $C_n + N_2$ by adding a new vertex v and some additional edges connecting v to $C_n + N_2$. Let N_2 be composed of the vertices a and b. Since F is connected, we must have edges $v \sim a$ and $v \sim b$, and at least one edge $v \sim x$ where $x \in C_n$. Since the induced subgraph on the vertices of $C_n \cup \{v\}$ is connected, we must have exactly one edge connecting v to C_n.

Since necessity is immediate, the proposition is established. ∎

Lemma 11.3. *$\{(C, F) \colon C \approx C_m$ and $F \approx pP(m)\}$ is definable.*

Proof. The criteria is that $F \approx pP(m)$ if and only if $C_m + K_2 \prec F$, and if

$m = 3$, then $F \approx pM(3)$, or if
$m > 3$, F is connected, $K_3 \leq F$, and if $C_m \prec R \leq F$, then $R \approx C_m \rightarrow_1$ or $R \approx C_m + N_1$.

We only establish sufficiency.

Suppose F satisfies the conditions. Let F be constructed from $C_m + K_2$ by adding a new vertex v and possibly new edges incident with v. Let a and b be the two vertices of K_2. We may assume $m > 3$.

Let $m > 3$. Since F is connected, we have at least one edge $v \sim x$ for $x \in C_m$. Then we must have exactly one edge since the induced subgraph on $C_m \cup \{v\}$ is connected. Since $K_3 \leq F$, we must have edges $v \sim a$ and $v \sim b$. ∎

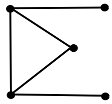

Figure 5: The Bull.

We describe a cover of P_4 formed by adding a new vertex to P_4 and exactly two new edges connecting the new vertex to the two vertices of P_4 which have degree 2. This graph is the *Bull*, and in light of Proposition 9.4 it is definable.

Here is how to encode a function $f \colon [n] \rightarrow [m]$. Define the graph $\sigma(n, f, m)$ over the vertex set

$$\sum_{i=1}^{n} C_{3+i} + \sum_{i=1}^{m} K_{3+i} + N_n,$$

with the following edge relations: let $\{v_1,\ldots,v_n\}$ be the vertices of N_n; choose vertices $x_i \in C_{3+i}$ and $u_j \in K_{3+j}$ for $i=1,\ldots,n$ and $j=1,\ldots,m$; take all the edges of $\sum_{i=1}^n C_{3+i} + \sum_{i=1}^m K_{3+i}$ together with edges $x_i \sim v_i$ for $i=1,\ldots,n$, and edges $v_i \sim u_{f(i)}$ for $i=1,\ldots,n$; these are the only edges.

Notice the choices of u_i and x_i are immaterial. Below is $\sigma(n,f,m)$ for $f(1)=f(2)=1$, $f(3)=2$.

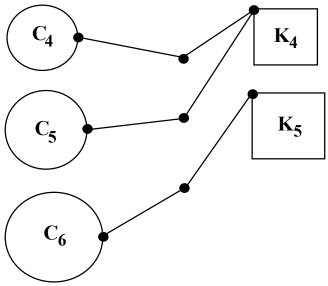

Figure 6: $\sigma(3,f,2)$.

Remark 11.4. It may be possible to replace the conditions listed in the next proposition by a more elegant set, but the advantage is that the proof of sufficiency is straightforward to verify.

Proposition 11.5. *We have the following*:

(1) $\sigma(n,f,m) \approx \sigma(n',f',m')$ *if and only if* $n=n'$, $m=m'$ *and* $f=f'$.

(2) $\{(C_n,K_m,F)\colon n,m>0, F\approx\sigma(n,f,m)$ *for some* $f\colon [n]\to[m]\}$ *is a definable relation.*

Proof. We tackle statement (1). Observe that $\sigma(n,f,m)$ contains a unique copy of $\sum_{i=1}^n C_{3+i}$ and of $\sum_{i=1}^m K_{3+i}$. This implies we must have $n=n'$ and $m=m'$. For each $i\in[n]$, C_{3+i} appears in exactly one pointed sum $K_{3+f(i)} +_p C_{3+i}$ in $\sigma(n,f,m)$, and thus, also in $\sigma(n',f',m')$. This implies $f(i)=f'(i)$ for each $i\in[n]$, and so, $f=f'$.

For the second statement, the claim is that $F\approx\sigma(n,f,m)$ for some $f\colon[n]\to[m]$ if and only if

(1) $\sum_{i=1}^n C_{3+i} + \sum_{i=1}^m K_{3+i} \prec_n F$.
(2) For $i\in[n]$, C_{3+i} is not a connected component of F.
(3) For $i\in[n]$, if $C_{3+i}\prec R\le F$, then $R\approx C_{3+i}\to_1$ or $R\approx C_{3+i}+N_1$.
(4) For $i,j\in[n]$, $i\ne j$, if $C_{3+i}+C_{3+j}\prec R\le F$, then $R\not\approx C_{3+i}+_p C_{3+j}$.
(5) For each $i\in[n]$, $C_{3+i}\to_2\le F$.
(6) For each $i,j\in[n]$, $i\ne j$, $\gamma(3+i,3+j)\not\le F$.
(7) For $j\in[m]$, if $K_{3+j}\prec R\le F$, then $R\approx K_{3+j}\to_1$ or $R\approx K_{3+i}+N_1$.
(8) For $i\in[n]$, $P(i)\not\le F$ and $pP(i)\not\le F$.
(9) **Bull** $\not\le F$.

Suppose F satisfies the conditions. Condition (1) implies we can construct F by adding the vertices $\{v_1,\ldots,v_n\}$ to $\sum_{i=1}^n C_{3+i}+\sum_{i=1}^m K_{3+i}$ and possibly some new edges incident with the vertices $\{v_1,\ldots,v_n\}$. Condition (2) implies each C_{3+i} is adjacent to at least one vertex of $\{v_1,\ldots,v_n\}$. Condition (3) implies that if C_{3+i} is adjacent to some v_k, then there is a unique vertex $x_i \in C_{3+i}$ such that $x_i \sim v_k$. Condition (4) implies no two distinct C_{3+i} and C_{3+j} are adjacent to the same vertex of $\{v_1,\ldots,v_n\}$, and therefore, no C_{3+i} can be adjacent to more than one vertex of $\{v_1,\ldots,v_n\}$. We may reorder the vertices of $\{v_1,\ldots,v_n\}$ so that for each $i\in[n]$, v_i is the unique vertex adjacent to C_{3+i} by a unique edge. Condition (6) implies $v_i \not\sim v_j$ for $i\neq j$.

Condition (5) implies each v_i is adjacent to a vertex of some K_{3+j}, and condition (8) implies it is not connected to any other vertex of K_{3+j}, nor to any vertex of K_{3+r} for $r\neq j$. For each $i\in[n]$, let $f(i)=j$ where K_{3+j} is adjacent to v_i. Let $u_{f(i)}\in K_{3+f(i)}$ such that $v_i\sim u_{f(i)}$. Condition (9) asserts that whenever $f(i)=f(j)$ for $i\neq j$, then $u_{f(i)}=u_{f(j)}$; that is, a unique vertex is chosen in $K_{3+f(i)}$ so that whenever v_k is connected to the clique $K_{3+f(i)}$, it is connected by that vertex. The map $f\colon[n]\to[m]$ is the function we are after, and altogether we have shown $F\approx\sigma(n,f,m)$. ∎

12. Second-Order Definablity

We define a small category $\mathcal{CG}$. The objects Obj$\mathcal{CG}$ are the graphs of $\mathcal{G}$ and the morphisms Mor$\mathcal{CG}$ are just the graph homomorphisms between graphs in $\mathcal{G}$. In order to discuss definable relations in the category, we need an appropriate first-order language. Formally, we can treat the category as a 2-sorted first-order structure, with one sort for objects and another sort for morphisms, together with a ternary relation over the sort of morphisms which reflects composition, and two unary functions dom, codom: Mor$\mathcal{CG}\to$ Obj$\mathcal{CG}$ which map morphisms to their domains and codomains. The category structure is then described by the standard category axioms in this 2-sorted first-order language.

Perhaps, it is more natural to treat the category less formally by taking as basic relations membership as an object $X\in$ Obj$\mathcal{CG}$, membership as a morphism $f\in\mathcal{CG}(X,Y)$ and composition as a binary map $\circ\colon\mathcal{CG}(X,Y)\times\mathcal{CG}(Y,Z)\to\mathcal{CG}(A,C)$. Since the basic relation $f\in\mathcal{CG}(A,B)$ explicitly references the domain and codomain, a morphism can also be written as a triple $F=(A,f,B)$ where $f\colon[n]\to[m]$ is a particular set map with $n=|A|$ and $m=|B|$. The first-order language of the category has existential and universal quantifiers ranging over objects, quantifiers ranging over morphisms in $\mathcal{CG}(X,Y)$, and is built up from composition using conjunction, disjunction, negation, and equality.

We enrich the small category by adding four new constants to the language and denote the resulting structure by $\mathcal{CG}'$. We add the constants $\mathbf{K}_2$, $\mathbf{P}_3$, and the two maps in $\mathcal{CG}(N_1,\mathbf{K}_2)=\{\mathbf{t},\mathbf{b}\}$. Notice that N_1 is a terminal object in $\mathcal{CG}'$, and so is definable. We will show in Theorem 12.3 that when restricted to the objects $\mathcal{G}$ of the category, $\mathcal{CG}'$ is equivalent to $\langle\mathcal{G},\leq,P_3\rangle$ in expressive power. The first step is to show that the isomorphic substructure relation is definable in $\mathcal{CG}'$.

A morphism $f\in\mathcal{CG}(A,B)$ is said to be a *monomorphism* if and only if for all $X\in$ Obj$\mathcal{CG}$ and for all $g,h\in\mathcal{CG}(X,A)$, $fg=fh\leftrightarrow g=h$. We say $f\in\mathcal{CG}(A,B)$ is an *epimorphism* if and only if for all $Y\in$ Obj$\mathcal{CG}$ and for all $g,h\in\mathcal{CG}(B,Y)$, $gf=hf\leftrightarrow$

$g = h$. The property of being a monomorphism or an epimorphism is by definition first-order definable in the language of the category. In general categories, we do not formally have access to the "inner" structure of the objects and don't expect to definably capture the property of injectivity or surjectivity; likewise, the property that $f \in \mathcal{CG}(A,B)$ is an embedding refers to the relational structure of A and B which is not included in the 2-sorted language of the category. In the case of graphs, injectivity happens to be equivalent to being a monomorphism and surjectivity is equivalent to being an epimorphism; thus, definable properties.

We can use the first-order structure of the category and the constants to definably manipulate the edge relations of object in the category. For any graph A and vertices $u, v \in A$, we see that $u \sim v$ in A if and only if where $x, y \in \mathcal{CG}(N_1, A)$ such that $x(0) = u$ and $y(0) = v$, then there exists $h \in \mathcal{CG}(\mathbf{K}_2, A)$ such that $h\mathbf{t} = x$ and $h\mathbf{b} = y$.

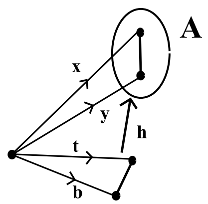

Figure 7: Reading the edge relation.

Now that we have a way to definably refer to the edge relation in objects, we can capture embeddings. We see $f \in \mathcal{CG}(A,B)$ is an embedding if and only if f is a monomorphism and whenever there exist $x, y \in \mathcal{CG}(N_1, A)$ and $q \in \mathcal{CG}(\mathbf{K}_2, B)$ such that $g\mathbf{t} = fx$ and $g\mathbf{b} = fy$, then there exists $h \in \mathcal{CG}(\mathbf{K}_2, A)$ such that $x = h\mathbf{t}$ and $y = h\mathbf{b}$. This means that the substructure relation of $\langle \mathcal{G}, \leq, P_3 \rangle$ is first-order definable in $\mathcal{CG}'$.

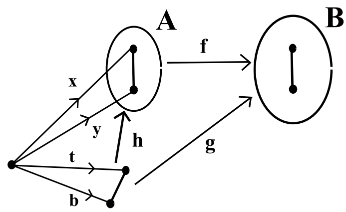

Figure 8: Capturing embeddings.

The next step is to define in $\langle \mathcal{G}, \leq, P_3 \rangle$ the basic relations of the category: $X \in \mathrm{Obj}\mathcal{CG}$, $f \in \mathcal{CG}(X,Y)$ and composition of morphisms. Recall, for any graph A, the o-presentation $P_n(A,B)$ contained an isomorphic copy $B \approx A$ on $[n]$ in such a way that the edge relation of B could be "read off" by $i \sim j$ in B if and only if $\gamma(n+2+i, n+2+j) \leq P_n(A,B)$. In this way, for any $A \in \mathcal{G}$, any graph isomorphic to $P_n(A,A)$ allows us to definably encode the edge relation of A in the structure $\langle \mathcal{G}, \leq, P_3 \rangle$.

Suppose we have graphs $G_i = \langle [m_i], r_i \rangle$ in the category and a morphism $F = (G_1, f, G_2)$ such that $f \colon [m_1] \to [m_2]$. We encode G_i as any graph isomorphic to $P_i = P_{m_i}(G_i, G_i)$ and encode F as any triple isomorphic to $M(F) = (P_1, \sigma(m_1, f, m_2), P_2)$.

In the next result, we see how to read off the values of a function f with statement (1), and how to capture that f is a homomorphism with statement (2).

Lemma 12.1. *We have the following*:

(1) *If $(U,V,W) \approx M(F)$ for $F = (G_1, f, G_2)$, then F (and f) are uniquely determined and for all $i \in [m_1]$ and $j \in [m_2]$, we have that $f(i) = j$ if and only if $K_{3+j} +_p C_{3+i} \leq V$.*

(2) *$(U,V,W) \approx M(F)$ for some $F = (G_1, f, G_2) \in \mathcal{CG}(G_1, G_2)$ if and only if where $m_i = |G_i|$, we have $U \approx P_{m_1}(G_1, G_1)$, $W \approx P_{m_2}(G_2, G_2)$, and $V \approx \sigma(m_1, f, m_2)$ for some $f : [m_1] \to [m_2]$; and whenever we have $1 \leq i, i' \leq m_1$ and $1 \leq j, j' \leq m_2$, $j \neq j'$, and $K_{3+j} +_p C_{3+i} \leq V$ and $K_{3+j'} +_p C_{3+i'} \leq V$, then $\gamma(m_1 + 2 + i, m_1 + 2 + i') \leq U$ implies $\gamma(m_2 + 2 + j, m_2 + 2 + j') \leq W$.*

Proof. For part (1), the first part of Proposition 11.5 and the second part of Proposition 9.8 guarantee that $(U,V,W) \approx M(F)$ if and only if $U \approx P_{m_1}(G_1, G_1)$, $W \approx P_{m_2}(G_2, G_2)$, and $V \approx \sigma(m_1, f, m_2)$. That $f(i) = j$ precisely when $K_{3+j} +_p C_{3+i} \leq V$ is explicit by construction.

For part (2), recall in the proof of Proposition 9.4 that $\gamma(m_1 + 2 + i, m_1 + 2 + i') \leq P_{m_1}(B_1, B_1)$ if and only if $i \sim i'$ in B_1. ∎

We now account for the composition of morphisms.

Lemma 12.2. *Let $F = (G_1, f, G_2)$ and $H = (G_2, h, G_3)$ with $|G_i| = m_i$ for $i = 1, 2, 3$. Let $M(F) \approx (P_1, \sigma_1, P_2)$ and $M(H) \approx (P_2, \sigma_2, P_3)$. Then $M(HF) \approx (P_1, \sigma_3, P_3)$ if and only if $\sigma_3 \approx \sigma(m_1, p, m_3)$ for some p such that: for all $i \in [m_1]$, $j \in [m_2]$, $k \in [m_3]$, we have that $K_{3+j} +_p C_{3+i} \leq \sigma_1$ and $K_{3+k} +_p C_{3+j} \leq \sigma_2$ imply $K_{3+k} +_p C_{3+i} \leq \sigma_3$.*

Let R be an isomorphism invariant N-ary relation over $\mathcal{G}$. We have already seen that if R is definable in $\langle \mathcal{G}, \leq, P_3 \rangle$, then it is definable in $\mathcal{CG}'$ by defining the isomorphic substructure relation. Suppose R is definable in $\mathcal{CG}'$; that is, there exists a formula $\Phi(x_1, \ldots, x_N)$ in the language of $\mathcal{CG}'$ such that

$$R = \left\{ (A_1, \ldots, A_N) \in \mathcal{G}^N : \mathcal{CG}' \models \Phi(A_1, \ldots, A_N) \right\}.$$

We need a formula $\Psi(x_1, \ldots, x_N)$ in the language of $\langle \mathcal{G}, \leq, P_3 \rangle$ such that

$$R = \left\{ (G_1, \ldots, G_N) \in \mathcal{G}^N : \ \langle \mathcal{G}, \leq, P_3 \rangle \models \Psi(G_1, \ldots, G_N) \right\}.$$

The strategy is to define, by induction on the complexity of $\Phi(x_1, \ldots, x_N)$, a formula $\hat{\Phi}(x_1, \ldots, x_N)$ so that whenever $A_i \approx B_i$ with $A_i \in \mathcal{G}$, and $|A_i| = k_i$ for $i = 1, \ldots, N$ we have

$$\mathcal{CG}' \models \Phi(B_1, \ldots, B_N) \quad \text{iff} \quad \langle \mathcal{G}, \leq, P_3 \rangle \models \hat{\Phi}(P_{k_1}(A_1, B_1), \ldots, P_{k_N}(A_N, B_N)).$$

Since R is isomorphism invariant, we can then take $\Psi(x_1, \ldots, x_N)$ to be

$$(\exists u_1, \ldots, u_N)\big(\hat{\Phi}(u_1, \ldots, u_N) \wedge (\text{“there exist } v_i \text{ such that } k_i = |x_i| \text{ and}$$

$$u_i \approx P_{k_i}(x_i, v_i) \text{ for } i = 1, \ldots, N\text{”})\big).$$

We have kept the notation and terminology of o-presentations and map encodings in strict faith with [9]. This guarantees our analogue of ([9, Theorem 3.8]) for graphs below goes through with the exact same translation scheme.

Theorem 12.3. *Let R be an isomorphism invariant relation over $\mathcal{G}$. Then R is first-order definable over $\langle \mathcal{G}, \leq, P_3 \rangle$ if and only if it is first-order definable in the language of $\mathcal{CG}'$.*

The first-order language of $\mathcal{CG}'$ is strong enough to capture second-order properties when restricted to $\mathcal{G}$. To see this, we definably build inside the category a set of isomorphic copies of the graphs in $\mathcal{G}$ so that definable relations in the language $\mathcal{CG}'$ can take the place of second-order concepts applied to the copies.

Using the bijection $G \ni u \leftrightarrow x \in \mathcal{CG}(N_1, G)$ such that $x(0) = u$, we see that given any $G \in \mathcal{G}$, we can construct an isomorphic graph

$$\hat{G} = \langle \mathcal{CG}(N_1, G), \hat{r} \subseteq \mathcal{CG}(N_1, G) \times \mathcal{CG}(N_1, G) \rangle \approx G, \tag{12.1}$$

where both the set of vertices and the edge relation $\hat{r}$ have first-order definitions in the language of the small category $\mathcal{CG}'$. It is not difficult to see that the set of such graphs $\{\hat{G} \colon G \in \mathcal{G}\}$ is definable.

Following the procedure outlined in [9, Sec. 3.1], we can use the first-order language of the category applied to the structures in $\{\hat{G} \colon G \in \mathcal{G}\}$ to parametrize arbitrary subsets of finitary cartesian products. To see this, take $\hat{G}_1, \ldots, \hat{G}_m$ and $\hat{R}$ a subset of the cartesian product of their universes; that is, $\hat{R} \subseteq \mathcal{CG}(N_1, G_1) \times \cdots \times \mathcal{CG}(N_1, G_m)$. By the bijection in the previous paragraph there is a corresponding relation $R \subseteq G_1 \times \cdots \times G_m$. If $|R| = k$, we shall use the maps in $\mathcal{CG}(N_1, N_k)$ to parametrize k-element subsets of the product in the same way as the maps of $\mathcal{CG}(N_1, A)$ parametrize the elements of A. If $\pi_i \colon G_1 \times \cdots \times G_m \to G_i$ denotes the i-th projection of the cartesian product as sets, then for any fixed bijection $p \colon [k] \to R$ there is a fixed sequence of morphisms $p_i \in \mathcal{CG}(N_k, G_i)$ given by $p_i = \pi_i \circ p$. This follows since any set map $\alpha \colon [k] \to G$ corresponds exactly to a morphism of the trivial graph N_k into G. An arbitrary tuple in R is then specified by $(p_1(s), \ldots, p_m(s))$ where $s \in [k]$. With this choice of $(p_1, \ldots, p_m)$, we have the description $\hat{R} = \{(q_1, \ldots, q_m) \in \mathcal{CG}(N_1, G_1) \times \cdots \times \mathcal{CG}(N_1, G_m) \colon q_i = p_i \circ q$ for some $q \in \mathcal{CG}(N_1, N_k)\}$.

In this way, the first-order language of $\mathcal{CG}'$ restricted to the structures $\{\hat{G} : G \in \mathcal{G}\}$ is equivalent to a second-order language which has variables ranging over the elements of $\hat{G}$, variables for the morphisms between objects, and can express the edge relation in objects, application of morphisms to elements, composition of morphisms, and equality of elements and morphisms, and the apparatus to quantify over arbitrary subsets of finite products. Altogether, using the isomorphism in Equation (12.1), we see that the first-order language of $\mathcal{CG}'$ when restricted to the objects of the category is equivalent in expressive power to a full second-order language of simple graphs over the same set of objects. Theorem 12.3 then yields the following:

Theorem 12.4. *For every sentence ϕ in the second-order language of simple graphs, there is a formula $\Phi(x, y)$ in the first-order language of the quasi-ordered set $\langle \mathcal{G}, \leq \rangle$ such that a graph A in $\mathcal{G}$ models ϕ if and only if $\langle \mathcal{G}, \leq \rangle \vDash \Phi(A, P_3)$.*

It then follows that for graphs, the class of finite models of a second-order sentence is closed under edge-complementation if and only if the set of isomorphic copies in $\mathcal{G}$ are definable without constants in the substructure relation $\leq$.

Example 12.5. Perfect graphs are definable. According to [2], finite perfect graphs omit cycles of odd length greater than 3 and their complements. By the comments after Proposition 8.1, we can definably accomplish addition with the cardinality of cycles; therefore, it is easy to see that the set of odd cycles of length greater than 3 are definable. Then the set of finite perfect graphs is definable without constants.

Let $\langle N_{>0}, +, \times \rangle$ denote the structure over the set of positive integers such that the operations of addition and multiplication have their usual meaning. Clearly, Proposition 8.1 and the previous result allow us to define the ternary relations of addition and multiplication over the set of cliques, and so establish a first-order interpretation of $\langle N_{>0}, +, \times \rangle$ into $\langle \mathcal{PG}, \leq, P_3/\approx \rangle$. According to [18, Theorems 7 and 10], we have the following.

Corollary 12.6. *The elementary theory of $\langle \mathcal{PG}, \leq \rangle$ is undecidable and not finitely axiomatizable.*

Example 12.7. Comparability graphs are definable. Using Theorem 12.4 it suffices to find a second-order axiomatization, and since in second-order logic we can quantify over relations, we can directly use the definition: a simple graph $G = \langle V, E \rangle$ is a comparability graph if and only if there exists R such that $\langle V, R \rangle$ is an ordered set and for $u \neq v$, $(u, v) \in E$ iff $(u, v) \in R$.

Example 12.8. By Theorem 12.4, planar graphs are definable. To see this we will use the characterization of finite planar graphs in [16]. For a graph $G = \langle V, E \rangle$, the incidence poset is the height two ordered set $P(G) = \langle V \cup E, \leq \rangle$ where the only non-trivial relations are of the form $u \leq (v, w)$ provided $u = v$ or $u = w$. For a poset $P = \langle A, \leq \rangle$, a linear extension of P is a linear order $L = \langle A, \leq_1 \rangle$ such that $\leq \subseteq \leq_1$. The *order dimension* of P is the smallest integer t such that $\leq$ is equal to the intersection of t linear extensions. The main result in [16] states that a finite graph is planar if and only if the order dimension of its incidence poset is at most 3.

Since the elements of the incidence poset $P(G)$ is comprised of two distinct types, vertices and edges, we will model linear extensions of $P(G)$ by certain binary, ternary and 4-ary relations which taken together essentially partitions the order relation of a linear extension by how it relates vertices and edges together. For a finite graph $G = \langle V, E \rangle$, the incidence relation $I_G \subseteq V^3$ is defined as $(u, v, w) \in I_G$ if and only if $u = v$ or $u = w$, and $(v, w) \in E$. An extension of I_G is a sequence of relations $L = (L_1, L_2, L_3, L_4)$ with $L_1 \subseteq V^2$, $L_2, L_3 \subseteq V^3$, and $L_4 \subseteq V^4$ such that $I_G \subseteq L_2$ and

- $(u, v, w) \in L_2$ implies $u = v$ or $u = w$, and $(v, w) \in E$,
- $(u, v, w) \in L_3$ implies $u = w$ or $v = w$, and $(u, v) \in E$,
- $(u, v, w, z) \in L_4$ implies $(u, v) \in E$ and $(w, z) \in E$.

It is easy then to write sentences to say $L = (L_1, L_2, L_3, L_4)$ partitions a linear order over $V \cup E$ which respects the distinction between vertices and edges in the definition of an extension of I_G. For example, a part of transitivity would include, for all u, v, w, x, z, $(u, v, w) \in L_2$ and $(v, w, x, z) \in L_4$ implies $(u, x, z) \in L_2$.

A simple graph $G = \langle V, E \rangle$ is planar if and only if there exist relations I_G and $L = (L_1, L_2, L_3, L_4)$, $L' = (L'_1, L'_2, L'_3, L'_4)$, and $L'' = (L''_1, L''_2, L''_3, L''_4)$ not necessarily distinct such that

(1) I_G is the incidence relation of G, and
(2) L, L' and L'' are extension of I_G which partition a linear order, and
(3) $I_G = L_1 \cap L_2' \cap L_3''$.

13. Concluding Remarks

In a review of substructure definability for ordered structures ([8–11]), and in extending these results to unordered structures of graphs, similarities in the general arguments abound, but in each case there is enough distinction for the development to start over. Having established strong definability results for these classes, can we abstract the combinatorial or model properties which may guarantee similar substructure definability in general universal classes?

In general, positive definability does not imply that the expressive power of the substructure relation captures second-order properties. For a simple example, we can consider the class of cliques $\mathcal{K}$. If $\mathcal{PK}$ denotes the set of finite isomorphism types, then $\langle \mathcal{PK}, \leq \rangle \approx \langle \mathbb{N} \leq \rangle$, and it is easy to see that a subset of $\langle \mathbb{N} \leq \rangle$ is definable precisely when it is finite or co-finite. But the cliques of even order are the finite models of a single second-order sentence.

For posets ([9]) and simple graphs ([9, Cor. 12.4]), the expressive power of first-order definability in the substructure relation is equivalent to modeling full second-order sentences when restricted to the finite members. We may consider a more restrictive language. For example, what fragment of the first-order language in the substructure relation corresponds to strictly first-order properties of finite graphs?

For tournaments, there is an obvious automorphism **rev** of the substructure ordering which comes from reversing the orientation of the edges. The counterexamples to the Reconstruction Conjecture for tournaments, and thus digraphs, discovered by Stockmeyer ([17]) appear in two infinite families (B_i, C_i) and (D_i, E_i). Interestingly, $\mathbf{rev}(B_i) = B_i$, $\mathbf{rev}(C_i) = C_i$, and $\mathbf{rev}(D_i) = E_i$. This is precisely what one must have if it is the case that the sets $\{B_i, C_i\}$ and $\{D_i, E_i\}$ are definable without constants. This prompts the following two questions.

Question 13.1. After adding a constant, is every finite isomorphism type of tournaments first-order definable in the poset of finite isomorphism types ordered by substructure, where **rev** is the only non-trivial automorphism? Do the universal theories of tournaments have positive definability?

Question 13.2. Is each pair of Stockmeyer's counterexamples $\{B_i, C_i\}$ and $\{D_i, E_i\}$ definable in the poset of finite isomorphism types ordered by substructure without adding a constant to the language?

We may depart from substructures, and instead define $A \preceq B$ if and only if A is a graph minor of B. In a natural way, form the poset $\langle \mathcal{PG}, \preceq \rangle$ of finite isomorphism types ordered by graph minor. What is the scope of definability in this partially ordered set? Does graph minor capture full second-order properties in the finite?

Acknowledgment. The author expresses his sincere gratitude to Ralph McKenzie for his suggestion of the first-order conditions in Proposition 5.3. This provided for a rather simple way to

capture connected graphs, and greatly simplified the development of the remaining constructions.

References

1. Burris, S., Sankappanavar, H.P.: A Course in Universal Algebra. Graduate Texts in Mathematics, Vol. 78. Springer-Verlag, New York (1981)
2. Chudnovsky, M., Robertson, N., Seymour, P., Thomas, R.: The strong perfect graph theorem. Ann. of Math. (2) 164, 51–229 (2006)
3. Harary, F., Holzmann, C.: Line graphs and Bipartite graphs. Rev. Soc. Mat. Chile 1, 19–22 (1974)
4. Ježek, J.: The lattice of equational theories. Part I: modular elements. Czech Math. J. 31, 127–152 (1981)
5. Ježek, J.: The lattice of equational theories. Part II: the lattice of full sets of terms. Czech Math. J. 32, 573–603 (1981)
6. Ježek, J.: The lattice of equational theories. Part III: definability and automorphisms. Czech Math. J. 3, 129–164 (1982)
7. Ježek, J.: The lattice of equational theories. Part IV: equational theories of finite algebras. Czech Math. J. 36, 331–341 (1986)
8. Ježek, J., Mckenzie, R.: Definability in Substructure Orderings I: finite semilattices. Algebra Univers. 61, 59–75 (2009)
9. Ježek, J., Mckenzie, R.: Definability in Substructure Orderings II: finite ordered sets. Order 27, 115–145 (2010)
10. Ježek, J., Mckenzie, R.: Definability in substructure orderings III: finite distributive lattices. Algebra Univers. 61, 283–300 (2009)
11. Ježek, J., Mckenzie, R.: Definability in substructure orderings IV: finite lattices. Algebra Univers. 61, 301–312 (2009)
12. Ježek, J., Mckenzie, R.: Definability in the lattice of equational theories of semigroups, I. Semigroup Forum 46, 199–245 (1993)
13. Kaddour, H.S., Pouzet, M., Trotignon, N.,: Claw-freeness, 3-homogeneous subsets of a graph and a reconstruction problem. Contrib. Discrete Math. 6(1), 86–97 (2011)
14. Mal'cev, A.I.: Universally axiomatizable subclasses of locally finite classes of models. Sibirsk. Mat. Ž 8(3), 1005–1014 (1967)
15. McKenzie, R.: Definability in lattices of equational theories. Ann. Math. Logic 3, 197–237 (1971)
16. Schnyder, W.: Planar graphs and poset dimension. Order 5, 323–343 (1989)
17. Stockmeyer, P.K.: The falsity of the reconstruction conjecture for tournaments. J. Graph Theory 1(1), 19–25 (1977)
18. Tarski, A., Mostowski, A., Robinson, R.: Undecidable Theories. North-Holland, Amsterdam (1953)