Forbidden Substructures and Combinatorial Dichotomies: WQO and Universality

Gregory Cherlin[1]

Department of Mathematics, Rutgers University
110 Frelinghuysen Road
Piscataway, N.J. 08854

Abstract

We discuss two combinatorial problems concerning classes of finite or countable structures of combinatorial type. We consider classes determined by a finite set of finite constraints (forbidden substructures). Questions about such classes of structures are naturally viewed as algorithmic decision problems, taking the finite set of constraints as the input. While the two problems we consider have been studied in a number of natural contexts, it remains far from clear whether they are decidable in their general form. This broad question leads to a number of more concrete problems. We discuss twelve open problems of varying levels of concreteness, and we point to the "Hairy Ball Problem" as a particularly concrete problem which we state both in direct model theoretic terms, and decoded as a completely explicit graph theoretic problem.

Keywords: universal graph, well quasi-order, antichain, model theory, decision problems
2000 MSC: 05C75, 03C13, 03C15, 03D35, 05C63

Contents

[1] Supported by NSF Grant DMS-0600940

1. Introduction

1.1. Dichotomies for Combinatorial Structures

We will discuss two problems which concern classes of combinatorial structures—in the first case finite structures, and in the second case countably infinite ones. The classes we consider are defined by finitely many constraints provided by "forbidden substructures."[2] Influenced by logic—complexity theory on the one hand, model theory on the other— we tend to put these problems in a very broad context, but open questions abound at all levels. A considerable body of concrete work has been undertaken on both problems in a number of contexts, but there is a great deal of similar territory remaining largely unexplored. Our survey includes some new results that we find clarifying. We have put most of the detailed discussion of the new results in three Appendices, referring to them as needed in the text, with an indication of the line of argument. This includes some results to the effect that "Here there be tygers," which are intended to justify some of the restrictions we impose.

One of the aims of general model theory has been to prove a dichotomy for the behavior of the most general classes of structures: the so-called "structure/nonstructure" alternative, in Shelah's parlance. According to this dichotomy, when one looks at large infinite models of first order theories, one either has a coherent structure theory which in the first instance allows one to estimate the number of models, and to proceed from there to more delicate results, or on the other hand one finds a degree of chaos which can be expressed in a number of ways, the essential point being that the behavior of the models in the nonstructured case is more a matter of set theory than of algebraic structure.

Are there any similar phenomena in the world of finite (or nearly finite) combinatorics? We will confine ourselves to classes of structures with very simple definitions, namely with classes defined by finitely many constraints of the simplest kind: forbidden substructures. We consider notions of "tameness" and "wildness" appropriate to this context, and we undertake to analyze the gap between the tame and the wild.

[2] Graph theorists and model theorists use the term "substructure" in distinct ways: see Note 2 in §4. We follow graph theoretic usage here.

The two notions of tameness with which we will work are the following: first, well-quasi-order; secondly, the existence of a countable universal object. If we followed the pattern of model theory exactly, we would be looking to show that the wild case is extremely wild in some sense; in the first of our two cases we doubt this, and in the second case, while it seems to be true, it is not really the point. For us, the natural question at this level is whether the separation between the tame and wild cases is *effective* (algorithmically decidable). Indeed, that is simply a precise way of stating that the two cases can be clearly separated. For our two interpretations of tameness—and no doubt, many others—it is completely unclear at this stage whether such a separation occurs. All one can really say to date is that when one works on instances of these problems, they seem difficult, and not entirely unlike some known undecidable problems.

Let us take up these two problems one at a time.

1.2. The WQO Problem

Here we deal with the class $\mathcal{Q}$ of all finite structures of a particular combinatorial type. This may be the class of (finite) graphs, tournaments, digraphs, permutation patterns, matroids, and such. We take a finite subset $\mathcal{C}$ of $\mathcal{Q}$, the *forbidden substructures*, and consider the subclass $\mathcal{Q}_{\mathcal{C}}$ of structures in $\mathcal{Q}$ containing no substructure isomorphic to any C in $\mathcal{C}$. A note on terminology: we use the term "substructure" here in much the same way that graph theorists use the term "subgraph;" and this is not consistent with standard model theoretic terminology. See Note 2, §4.3 for more on this point, and also §1.6 and Appendix 3 (§7).

As $\mathcal{Q}$ is not actually a set, one may prefer to cut it down by taking all structures under consideration to have their elements in a fixed countable set; or indeed by working with isomorphism types rather than structures. We will not concern ourselves with the choice of formalism.

The relation that interests us here is the embeddability relation on $\mathcal{Q}$: $a \leq b$ if a is isomorphic with a substructure of b. Then $\mathcal{Q}$ is a quasi-order, and the equivalence relation given by $a \leq b \leq a$ is the relation of isomorphism. All of these quasi-orders are *well-founded*, that is there is no infinite strictly descending sequence $a_1 > a_2 > \dots$.

In general, a quasi-order is said to be well-quasi-ordered (wqo) if it is both well-founded and contains no infinite antichain (i.e., set of pairwise incomparable elements). The problem we wish to consider—in its first formulation—is the following.

Problem (A)**.** *With $\mathcal{Q}$ and $\mathcal{C}$ specified, is $\mathcal{Q}_{\mathcal{C}}$ wqo? In other words, does $\mathcal{Q}_{\mathcal{C}}$ contain an infinite antichain?*

We consider some illustrative examples.

Fact 1.1.

1. *Let L be a finite linear tournament. Then the L-free tournaments are wqo (in fact of bounded size, by Ramsey's theorem) . . .*

2. *But if T is a nonlinear tournament, with at least 7 vertices, then the T-free tournaments are not wqo (by [33], because of two very special antichains serving to witness this in all cases).*

This gives the following corollary.

Corollary 1.2. *The finite tournaments T for which the class of T-free tournaments is wqo can be recognized in polynomial time.*

Results of this kind often have a paradoxical quality: Fact 1.1 doesn't actually tell us how to determine which side of the fence a particular constraint T will actually fall, if T is nonlinear and very small, nor does it give us any hint as to how one should find out in such cases. But once the number of cases left unsettled is finite, and the others are cleanly handled, the problem becomes polynomial time decidable. At the same time, it is precisely the finitely many cases left over that tend to be the real challenges in practice, and in the present instance it took extensive structural analyses of the classes $\mathcal{Q}_T$ associated with two of these "left over" tournaments T, and then an application of Kruskal's tree theorem [31], to convert this abstract statement into a definite answer.

Thus a proof that a problem is solvable is not at all the same thing as a solution, and the distinction is worth bearing in mind. But we find the question, whether such combinatorial problems are solvable in principle at a systematic level, to be one with its own interest.

At the level of generality of the problems we consider, algorithmic decidability per se is the natural question. But one curious feature of the wqo problem is that decidability results are obtained by noneffective methods, and that the resulting algorithms whose existence is proved are "good" in the conventional sense of polynomial time computability, even though no single correct algorithm is produced, and for that matter in certain cases no explicit bound on the degree of the associated polynomial can be extracted from the decidability proof. This is not a new phenomenon; it comes with the general territory of wqo theory [18, §8].

We restate our problem in the form that actually concerns us.

Problem $(A_{\mathcal{Q}})$**.** *With $\mathcal{Q}$ fixed, for example the class of finite tournaments, and with $\mathcal{C}$ varying, is Problem A effectively solvable (and if so, in polynomial time)? That is, is the function taking us from the specification of $\mathcal{C}$ to the answer, a computable function?*

Our thesis is this: if the classification of classes of the form $\mathcal{Q}_{\mathcal{C}}$ into wqo and non-wqo cases can be accomplished effectively, then we have a real dichotomy, with a gap between the two cases (possibly revealed by the proof!); and if not, then this expresses the absence of any clear borderline separating the two alternatives.

This puts us in mind of Wang's *domino problem*: given a finite set of tile types (square tiles of fixed size, but with any of a finite number of "colors"), and some tiling rules allowing only certain pairs of tile types to be juxtaposed horizontally, and certain pairs vertically, to determine whether the plane can be tiled completely using tiles of the specified types, and respecting the constraints. (One can also encode the "colors" by small variations in shape, so that this becomes literally a geometric problem of completely covering the plane.) This problem was shown to be undecidable by Berger [2]. In particular, this refuted a conjecture by Wang that any set which allows such a tiling would allow a periodic one.

Wang's conjecture would have given a clean separation between the two possibilities; and we think Berger's result may reasonably be taken to mean that there is no clear line of separation, in the case of Wang's problem.

And we raise the same question for the wqo problem.

1.3. Universal graphs with forbidden subgraphs

To approach our second problem, we first consider some examples.

Fact 1.3.

1. *[45, 1964] There is a universal countable graph.*
2. *[28, 1988] For any path P, there is a universal countable P-free graph.*
3. *[20, 1981] There is no universal countable C_4-free graph.*

Universality of G usually means the following here: for each countable graph H in the class under consideration, there is an *induced subgraph* of G isomorphic to H. If we require only an embedding as a subgraph, we speak of *weak universality*, and on such occasions, we may refer to our usual notion of universality as *strong universality*, for emphasis. One prefers to prove the existence of universal graphs in the strong form, and the nonexistence in the weak form, taking special note of the rare instances where a weakly universal graph exists but a strongly universal one does not.

The *Rado graph* [45] is often built probabilistically, or explicitly, though neither approach lends itself well to the natural generalizations.

The universal graph with a *forbidden path* [28] is handled differently: this is based on a structure theorem for the class in question.

Lastly, with the cycle C_4 forbidden, we find ourselves on the other side of the fence. We will describe how one obtains negative results in such cases much later, in §3.

Now we consider the general case, specifying a class $\mathcal{Q}$ of finite or countably infinite combinatorial structures, and a finite constraint set $\mathcal{C}$, whose elements are finite, and preferably connected as well. The appropriate notion of connectedness for structures of general type is connectedness of the *Gaifman graph,* whose vertices are the elements of the structure, with edges between any two vertices which occur within some n-tuple for which one of the basic relations of the structure holds.

Problem (B)**.** *With $\mathcal{Q}$ and $\mathcal{C}$ specified, determine whether there is a universal structure in $\mathcal{Q}_{\mathcal{C}}$.*

Of course we mean to ask this in a more algorithmic spirit, as follows.

Problem $(B_{\mathcal{Q}})$**.** *With $\mathcal{Q}$ fixed, for example the class of countable graphs, and with $\mathcal{C}$ varying, is Decision Problem B effectively solvable? That is, is the function taking us from the specification of $\mathcal{C}$ to the answer, a computable function?*

This problem remains open, and probably quite difficult, even in the case of a single constraint C. But in view of the more recent developments in the area, which we will get into in §3, I am convinced that at least in the case of a single constraint, this should be a decidable problem.

Beyond that, I see no strong reason to conjecture what will happen in general. Here again, the domino problem comes to mind. We shall see something more about the relationship between the two problems at the end of §3 and in Appendix 3 (§7).

Model theorists should again take note of the terminological point that forbidden *subgraphs* are not forbidden *induced subgraphs*. We are using the customary terminology of graph theory here, but when we move to a broader context we will be dealing with the parallel notions of "substructures" and "induced substructures" rather than "subgraph" and "induced subgraphs".

1.4. Universality with one constraint

We will not undertake a discussion of the theory underlying the analysis of universality problems till §3, but we do want to say more at this stage about the case of one constraint.

A good deal of the evidence for the decidability of that case is found in the proof of the following.

Fact 1.4.

1. *[17] If C is a finite 2-connected graph, then there is a universal countable C-free graph if and only if C is complete.*

2. *[10] If C is a tree, then there is a universal countable C-free graph if and only if C is either a path, or derived from a path by adjunction of a single edge.*

It seems that the general case of a single constraint may behave like an amalgam of these two special cases, and that some elements of a general proof are in hand. Any connected graph C can be viewed as built up from its *blocks* (maximal 2-connected subgraphs) along a tree. We conjecture that a necessary condition for existence of a universal C-free graph is that the blocks should be complete, and there is some theoretical basis for this, given in §3. We do not expect the tree structure to be as simple as the foregoing fact might suggest, but we do think the underlying tree structure will be severely limited. If that fails, then all bets are off—and there are some basic case studies that still need to be carried out.

In general there will have to be some nontrivial interaction between the sizes of blocks and the structure of the tree of blocks. For example, consider 2-bouquets $K_m \wedge K_n$: these are formed from two complete graphs K_m and K_n by joining them at one common vertex, making something perhaps more like a bow-tie than a bouquet.

Fact 1.5 ([14]). *Let $C = K_m \wedge K_n$ be a 2-bouquet of type (m, n). Then there is a universal countable C-free graph if and only if the following two conditions are satisfied:*

1. $\min(m, n) \leq 5$;
2. $(m, n) \neq (5, 5)$.

In terms of the underlying tree structure, we are considering here a path of length 2. I would have predicted an answer of a slightly different type, more in the vein of:

$$\min(m, n) \leq 4$$

(or even 3)—and without the "discontinuity" at $(5, 5)$.

A considerable amount of computation goes into results of this type. But we have a systematic theory telling us what sort of computation is relevant, and this will be discussed in §3. One case where we are at a loss to complete this computation will be proposed to the reader as Problem 5.

What really gives us optimism about the general case of a single constraint is the style of the proofs in Fact 1.4. The Füredi-Komjáth argument is certainly malleable enough to cover more than the case of a single block (and a more general form is given in the original paper). And the tree analysis is based on a simple notion introduced by Shelah, called *pruning*, which in the case of trees brings us quickly down to the consideration of a finite number of topological trees (though not a finite number of isomorphism types as graphs). Fortunately, pruning is not restricted to trees—we can use the underlying tree structure of the blocks. We will come back to this as well in §3.

1.5. Tame or wild

The wqo/not-wqo distinction is a natural enough way to make a "tame vs. wild" distinction. The same cannot be said for the question of existence of a universal graph, at first glance. Still, we shall see that the latter also translates into a comprehensible tameness condition.

On the wqo side, it is agreeable that once a class becomes tame, so do its subclasses. On the universality side, we begin with Rado's graph in the class of all graphs, so we start on the tame side, then shrink to various wild classes, and then back again to tame cases. Or starting with a large 2-bouquet of type (m, n), we can shrink to $(5, n)$, then to $(5, 5)$, and finally $(5, 4)$ and wander in and out of the tame side. The virtue of "pruning" will be that it provides an antidote to this malaise: if one lops off blocks from the constraint (with sufficient uniformity) then this sort of erratic behavior will be eliminated.

But the theory presented in §3 provides a different and more useful perspective. There we will see that the class $\mathcal{C}$ of constraints naturally produces a notion of "algebraic closure," which we will denote by $\mathrm{acl}_{\mathcal{C}}(A)$, and that the essential question is the behavior of this operator. In the favorable case, the algebraic closure of a finite set is finite, and then there will be a universal object. In the unfavorable case, where $\mathrm{acl}_{\mathcal{C}}(A)$ is infinite with A finite, we cannot immediately conclude that there is no universal $\mathcal{C}$-free graph. But apart from a few simple cases with a particularly straightforward structure, this has been the case in practice.

So the tameness in question here is the tameness of an associated closure operation, which in its most rudimentary incarnation (in locally finite graphs) is given by simply taking the connected components of the elements involved, but which in general is connected to the structure of the constraints in $\mathcal{C}$ in a subtle way. The task of the general theory is to lay out this connection, and then the bulk of the concrete results come from understanding something about this closure operator in practice.

1.6. Varying $\mathcal{Q}$

Our discussion so far has ignored the effect of varying the context $\mathcal{Q}$ in which the problem is treated. Decision Problem A arose in the context of tournaments, because it was associated with a natural decision problem relating to homogeneous digraphs in view of [8]; and Decision Problem B comes directly from the graph theoretic literature, e.g. [45, 41, 27, 28, 19, 26].

But these problems, and the theory that goes along with them, make perfectly good sense for combinatorial structures of arbitrary type. And that is the natural level at which to pose these problems as algorithmic questions. One may also look for "reduction theorems:" these would state that posing these problems in a single natural context exhausts the range of problems of this kind, up to effective reductions.

One reduction theorem has been proved to date: Decision Problem $B_{\mathcal{Q}}$, for general combinatorial structures, reduces to the same problem posed in the context of graphs with a coloring of the vertices by two colors [13]. This reduction was one step in a project aimed at proving the undecidability of the problem for graphs. The reduction theorem was intended to provide a key bridge, but this particular bridge lacks supports at either end. On the one hand, we never found a setting in which these problems could be shown to be undecidable; and on the other hand, we could not get that encoding to go into the class of graphs. Even so, the problem for graphs with colored vertices is amenable to the same range of techniques as the problem for graphs, and knowing that this is the general case does provide some additional incentive for taking it up in that form.

As far as Problem A_Q is concerned, the issue of reductions has never been taken up seriously. It is not clear how one would approach that, but it is worth looking into.

1.7. Plan of the Paper

Our plan for the rest of the paper is to take up the two problems discussed again from the point of view of the general theory, and to indicate how that theory has shaped the work to date and what it suggests about the natural continuation in concrete cases—in the absence of a major breakthrough from the side of undecidability, of which there is little sign at present.

§2 deals with Decision Problem A_Q. In the first two subsections we describe work by Latka [33, 34, 35] on this problem for the case of tournaments, in the case of a single forbidden subtournament, and a general finiteness result from [9] which amounts to the statement that results qualitatively similar to Latka's also hold for classes of tournaments defined by any *fixed* number of forbidden subtournaments. Our finiteness theorem provides remarkably little information about the actual content of such results, as the method of proof is a nonconstructive argument typical of the theory of well quasi-orders.

In the remaining parts of §2 we look at some other instances of the wqo problem. Guoli Ding [16] found that wqo problems for the case of (symmetric) graphs with forbidden subgraphs are very simple (§2.3). He showed that up to equivalence there are only two minimal antichains, and only one of these antichains is isolated. Turning to the case of permutation patterns (§2.4), we take note of considerable recent progress on the structure of minimal antichains and the wqo problem. As in the case of tournaments, there is as yet no complete classification ([52]). In §2.5 we examine a much simpler quasi-order (on vertex colored paths) in which one can identify the isolated antichains explicitly and solve Decision Problem A_Q. There are very natural embeddings (encodings) from the quasi-order of colored paths to the quasi-order of tournaments that with one small variation account for the known minimal antichains of tournaments, as discussed in §2.6. It is unclear to me whether the known antichains of permutations can be accounted for by encodings of ordered paths, or any similarly elementary combinatorial structures. There is a systematic theory of construction in that case, but it is more subtle than in the case of tournaments.

In §3 we take up Problem $B_{\mathcal{Q}}$ in the context of graphs, and we do not stray much from that context. Our survey focuses on the general theory of [11] and the applications in [17, 10, 14], leading to the suggestion that something like an effective solution should be available in the case of one constraint, giving in particular the decidability of the problem in that case. We remain completely uncertain whether the problem is decidable for an arbitrary finite set of constraints, and we take note of the attempt in [13] to build a machine for interpreting some undecidable problem. Our tools for solving the problem have gotten well ahead of our tools for proving undecidability, which may not say much about what the final result will be, but explains the current directions of research. In particular we will explain how Shelah's simple but powerful "pruning" operation allows something like an inductive approach to the problem [10].

Apart from these matters, which are covered in the literature, we address three others: (1) we give some motivation for the consideration of Problem $B_{\mathcal{Q}}$ in terms of forbidden substructures rather than in terms of forbidden *induced* substructures; (2) we treat the special case of Problem $B_{\mathcal{Q}}$ in which the graphs involved have bounded vertex degree; (3) we show that an important conjecture relating to universal graphs with one constraint (Conjecture 3) will fail in the context of an arbitrary finite set of constraints.

Concerning the first point, we show in Appendix 3 (§7) that Problem $B_{\mathcal{Q}}$ becomes undecidable if we allow a finite set of *induced subgraphs* to be forbidden. In §3.6 we give a weaker result with similar content and a more direct proof.

For the second point, we sketch a proof that Problem $B_{\mathcal{Q}}$ becomes decidable when restricted to graphs of bounded vertex degree (Proposition 3.4, with more details given in Appendix 2 (§6)). The general theory of [11] degenerates in this case to a straightforward study of connected components of $\mathcal{C}$-free graphs. In particular our basic Theorem 2 becomes obvious when specialized to that context. It may be useful to think of the theory in general as an extension of that analysis. But the solution to Problem $B_{\mathcal{Q}}$ for the case of bounded vertex degree does not suggest a similarly direct solution for the general problem, even with this theory in place.

The third point relates to the theory of algebraic closure introduced in §3.2. The complexity of this operator reflects the structure of the constraint set. As we shall see, Conjecture 3 of §3.3 would support a direct approach to decidability for universality problems in the case of one constraint. The fact that the natural extension of this conjecture to the case of a finite constraint set fails argues against such a direct approach in general.

The proof of Proposition 3.4 got a little out of hand, so we just sketch the proof in §3.4, and give more detail in Appendix 2. In both §3.4 and the appendix, we pay more attention to the underlying structural analysis than to the decidability question per se.

One of the more concrete conjectures in the present paper is the Hairy Ball Conjecture of §3.4, which we take some pains to make explicit in purely graph theoretic terms in §3.5. This concerns an infinite family of constraints C for which the existence of a universal C-free graph is plausible, though far from certain. Such families are rare, and in the past have been fairly easy to identify when they do exist. The obstacles to the analysis in this case appear to be essentially graph theoretic.

§4 concludes the paper with a review of open problems touched on in §§2,3, and concludes with some technical notes concerning matters that a reader—particularly, a reader with a background in model theory—might expect to see addressed somewhere.

The three concluding appendices alluded to then follow, with detailed discussions of three results which are discussed more briefly in the exposition: universality problems for graphs of bounded vertex degree, universal graphs for which the associated operation of algebraic closure is not unary, and the undecidability of universality problems when the constraints are forbidden induced subgraphs.

2. Minimal Antichains in Well-Founded Quasiorders

The quasiorders that concern us have the following two properties.

- They are *well-founded*: there is no infinite, strictly decreasing chain;
- They are essentially countable, and effectively (and even, *efficiently*) presented.

While we are concerned with arbitrary finite graphs, digraphs, and the like, and hence with proper classes of structures, this could all be rephrased in terms of isomorphism types; or with less need for paraphrasing, in terms of structures whose elements are taken from a fixed countable set. When we deal with questions of effectivity some such approach should be taken, but we leave the details aside.

Let $\mathcal{Q}$ be a well-founded and effectively presented quasiorder. For $C \subseteq \mathcal{Q}$, $\mathcal{Q}^{\geq C}$ is $\{q \in \mathcal{Q} : \exists c \in C, q \geq c\}$, and $\mathcal{Q}_C = \mathcal{Q} \setminus \mathcal{Q}^{\geq C}$, an ideal (i.e., lower set) of $\mathcal{Q}$. We normally take C finite. The problem that concerns us is the following.

Problem ($A_{\mathcal{Q}}$)**.** *Is there an effective (more particularly, polynomial-time) procedure to determine whether, for a given finite $C \subseteq \mathcal{Q}$, the ideal $\mathcal{Q}_C$ is wqo?*

For this question to be meaningful, at a minimum the elements of $\mathcal{Q}$ must be coded effectively; but for the question to be not only meaningful, but reasonable, the structure that $\mathcal{Q}$ imposes on those elements must also be given effectively.

Our aim is to present a finiteness theorem for Problem $\mathcal{A}_{\mathcal{Q}}$ in the context of well-founded quasiorders in general, which goes some distance toward solving the restricted version of Problem $A_{\mathcal{Q}}$ in which the constraint set C is not only finite, but is taken to have k elements, with k fixed in advance. As the failure of wqo is witnessed by infinite antichains, our finiteness theorem aims to show that only finitely many antichains are relevant, for fixed k. After that, what remains to make this effective would be to show that the antichains involved are themselves sufficiently effective, in a precise sense. We sidestep the question of finding the relevant set of antichains effectively by fixing k; to solve Problem $A_{\mathcal{Q}}$ would involve knowing not only that a suitable finite set of antichains exists for each k, but also giving a method to find such a set effectively. Above all the "halting problem", that is, deciding when the promised finite set has been completely enumerated, is very hard in practice.

After proving the finiteness theorem, we will illustrate its content in the context of tournaments, and also of graphs. One proceeds by looking for the promised finite set, and as long as one has not found it, in practice one knows how to keep looking effectively; once one has found it, proving that the search is over becomes particularly difficult. In the nontrivial cases studied to date, it is here that Kruskal's tree theorem comes in; this says, roughly, that finite trees carrying labels taken from a wqo set form a new wqo set with respect to an appropriate notion of embedding.

Or we may put the difficulty like this: it is easier to realize that an ideal $\mathcal{Q}_C$ is *not* wqo, when that is the case, than it is to realize that an ideal $\mathcal{Q}_C$

is wqo, when that is the case. We will see this concretely below.

2.1. The finiteness theorem

An *antichain* in $\mathcal{Q}$ is a subset I whose elements are pairwise incomparable. There is a natural quasiordering on infinite antichains defined as follows:

$$J \leq I \text{ iff: } \forall j \in J \, \exists i \in I \, (j \leq i)$$

And then $J < I$ if $J \leq I$ and for some $j \in J$ and $i \in I$ we have $j < i$.

We call an antichain I *minimal* if it is infinite, and if for any infinite antichain J with $J \leq I$, we have $J \subseteq I$; but here, by abuse of notation, we will say $J \subseteq I$ if each element j of J is equivalent to an element i of I in the sense that

$$j \leq i \leq j$$

The following easy lemma is a version of Nash-Williams' "minimal bad sequence" argument [40].

Lemma 2.1. *Let $\mathcal{Q}$ be well-founded and let I be an infinite antichain. Then there is a minimal antichain J with $J \leq I$.*

Proof. Choose elements j_i ($i = 0, 1, 2, \dots$) as follows. Writing $J_k = \{j_i : i < k\}$, we let j_i be a minimal element of the set

$$\{j : \exists J^i \leq I \text{ infinite } (J_i \cup \{j\} \subseteq J^i)\}$$

Then j_i is defined for all i, inductively, and $J = \bigcup J_i$ is an antichain with $J \leq I$. We claim that J is a minimal antichain.

If $J' \leq J$ is an infinite antichain then we may adjust J' so that if $j' \in J'$ is equivalent to $j_i \in J$ then in fact $j' = j_i$. This will lighten notation. We claim that $J' \subseteq J$.

As $J' \leq J$ we have for each $j' \in J'$ an element $j_i \in J$ so that $j' \leq j_i$, and more particularly, either $j' < j_i$ or $j' = j_i$. If for some such pair we have $j' < j_i$, and we take i minimal, then $J_i \cup J'$ is an antichain containing $J_i \cup \{j'\}$, and then the fact that $j' < j_i$ violates the choice of j_i. So in fact $J' \subseteq J$ as claimed. $\square$

The following observations will be important when we come to the proof of the finiteness theorem for Problem A.

Remark 2.2.

1. *If I is a minimal antichain, then the ideal $\mathcal{Q}^{<I}$ defined as $\{q \in \mathcal{Q} : \exists q' \in I\, q < q'\}$ is a wqo (immediate);*

2. *If $\mathcal{Q}_1, \ldots, \mathcal{Q}_n$ are wqo, then $\prod Q_i$, with the pointwise ordering, is a wqo (Higman [22]).*

Now we state the main result of [9].

Theorem 1 ([9])**.** *Let $\mathcal{Q}$ be a well-founded quasiorder, $k \geq 0$ fixed. Then there is a finite set Λ_k of infinite antichains, such that for any set $C \subseteq \mathcal{Q}$ with $|C| \leq k$, the following are equivalent:*

1. *$\mathcal{Q}_C$ is not wqo;*

2. *There is some $I \in \Lambda_k$ such that*

$$I \subseteq^* \mathcal{Q}_C$$

That is, there is $I_0 \subseteq I$ finite, so that $I \setminus I_0 \subseteq \mathcal{Q}_C$.

Proof. We proceed by induction on k. The induction goes by proving a little more: that the antichains $I \in \Lambda_k$ may be taken to be minimal antichains as we go along.

Given Λ_k, we will of course take $\Lambda_{k+1} \supseteq \Lambda_k$, and the main point is to throw away all constraint sets $|C|$ of size $k+1$ which can already be understood in terms of Λ_k, and to see what remains.

Given $C \subseteq \mathcal{Q}$ with $|C| = k+1$, and any subset C' of C of cardinality k, if $\mathcal{Q}_{C'}$ is wqo we are done, and we discard this case. So for each such subset C', we suppose there is a corresponding minimal antichain $I' \in \Lambda_k$, witnessing the failure of wqo for $\mathcal{Q}_{C'}$. If $I' \subseteq^* \mathcal{Q}_C$ then we are again done, as we have already put I' into Λ_{k+1}. So if $c \in C$ is the unique element not in C', we have

$$I' \setminus \mathcal{Q}^{\geq c} \text{ is finite}$$

and, in particular,

$$c \in \mathcal{Q}^{<I'}$$

So now enumerating $C = (c_1, \ldots, c_{k+1})$ and letting $C'_i = C \setminus \{c_i\}$. we conclude that there are $I_i \in \Lambda_k$ so that

$$C \in \prod_i \mathcal{Q}^{<I_i}$$

Now we have observed above that each of the quasiorders $\mathcal{Q}^{<I_i}$ is wqo, and hence their product is also wqo. In particular the subset S consisting of those constraints

$$C \in \prod_i \mathcal{Q}^{<I_i}$$

for which $\mathcal{Q}_C$ is not wqo, is also wqo. Thus the set S has only finitely many minimal elements. For each of these finitely many constraint sets C we may choose a minimal antichain I_C in $\mathcal{Q}_C$, and we do this for each choice of $k+1$ antichains I_i out of Λ_k. Let Λ_{k+1} consist of Λ_k together with each of these additional minimal antichains I_C, and we are done. $\square$

This argument can be turned into an iterative procedure which is in fact a reasonable approach to concrete instances of Problem A. Namely, one looks first for Λ_0, which will be empty if $\mathcal{Q}$ is wqo, and will consist of a single minimal antichain otherwise. Then one bootstraps along inductively as suggested by the analysis given in the proof. A certain number of constraint sets of cardinality $k+1$ are already handled by the set Λ_k, and the remaining ones vary over a wqo family of constraint sets C of cardinality $k+1$; the expectation is that some of these will allow infinite antichains, and if specific constraint sets are chosen judiciously they may even have unique minimal antichains. In practice one may find these antichains quickly, as long as one does not fall into the trap of looking for an antichain in a case where $\mathcal{Q}_C$ actually is wqo. At a certain point, one suspects that Λ_{k+1} has been properly identified, and then matters take a turn for the worse. Some explicit structure theorems are needed to prove wqo for the remaining cases, and as these are extremal cases, the structural analysis involved may be elaborate.

Already in the case $k=1$, in the case of tournaments, all of these phenomena are visible, or would have been visible if the general theory had been in place when the work was done [33].

This analysis can be pushed a little further, so we will develop the formal side a little farther before turning to concrete cases. In [9] we expressed this in terms of a topological space whose points are equivalence classes of minimal antichains under the following equivalence relation:

$$I \sim J \text{ iff } \mathcal{Q}^{<I} = \mathcal{Q}^{<J}$$

So one could just as well take the corresponding ideals $\mathcal{Q}^{<I}$ to be the points of the space. The basic sets are then given by finite constraint sets C, where the open set U_C corresponds to $\{I \subseteq \mathcal{Q} : I \subseteq^* \mathcal{Q}_C\}$.

One focuses in particular on the *isolated* points of this space, that is the minimal antichains which are uniquely picked out by a finite constraint (up to equivalence). Indeed, an isolated antichain associated with a constraint C of size k must be in Λ_k, and if the isolated points are dense then $\Lambda = \bigcup_k \Lambda_k$ can be taken to consist of exactly the isolated minimal antichains. When that is the case, the issue of effectivity is simply one of an effective description of the isolated minimal antichains. We require the following.

1. An enumeration of the finite constraint sets C which isolate a minimal antichain I_C;

2. An algorithm for recognizing $\mathcal{Q}^{<I_C}$ for such C, given C (uniformly).

Most of the fundamental questions remain open, notably that of the density of the isolated points in cases of interest.

A curious feature of the inductive approach in terms of the sets Λ_k is that for a fixed $\mathcal{Q}$ and k, if one has identified Λ_k, and if the antichains are sufficiently effective in the above sense (here the first condition falls away, as the set in question is finite), then Problem A is decidable, for the given $\mathcal{Q}$ and for the parameter $k+1$. This is a "bonus" that can be extracted from the proof. At the end of the proof we see that there are only finitely many constraint sets that remain to be understood in order to make the transition from Λ_k to Λ_{k+1}, and in particular to solve Problem A fully for constraint sets of size $k+1$ it suffices to know Λ_k and just this finite set of additional constraints. In other words, taking this finite set of additional constraints as additional input to an appropriate algorithm, there is in fact an effective solution to the problem.

It may be tempting at this point to try to bypass the Λ_k entirely and work with the finiteness of the relevant collection of constraint sets to get a soft proof of decidability of Problem A restricted to fixed size. This does not work, as knowledge of Λ_k, while superfluous at stage k, becomes relevant at stage $k+1$.

2.2. Problem A for tournaments

Now let us consider, more concretely, wqo problems for classes of tournaments. So in this subsection we will suppose

$$\mathcal{Q} \text{ is the quasiorder of finite tournaments under embedding}$$

Wqo problems for classes of tournaments arose in the first instance in connection with the classification of homogeneous directed graphs [8], and were closely studied by Brenda Latka, with the main results presented given in [33], which relies on two substantial classification results worked out separately as [34, 35].

For $k = 0$, it seems the first published construction of an antichain in $\mathcal{Q}$ is due to Henson [21], who applied it to the construction of uncountably many homogeneous directed graphs. Henson's antichain is indeed minimal, and therefore it may serve as the unique element of Λ_0, though it turns out in retrospect not to be the optimal starting point. In any case, with this in hand, and looking, for $k = 1$, at constraints not settled by that example, one finds out that Λ_1 requires only two antichains, and that both are isolated, and quite straightforwardly effective. One also notices that with these two antichains in hand, one no longer needs Henson's originally antichain, which can now be discarded, though it will be needed subsequently as one of the antichains in Λ_2.

As we have indicated earlier, the correct identification of Λ_1 is important, but is only half the story. In this particular case, having this candidate for Λ_1 in hand already tells us the following, even before we confront the issue of its completeness:

For any nonlinear tournament C with at least
7 vertices, the ideal $\mathcal{Q}_C$ is not wqo.

—And one of the two minimal antichains found will serve as witness.

This looks like more than half the battle: all but finitely many cases have been disposed of, and we can show at this point that *if our conjecture for the set* Λ_1 *is correct,* it will suffice to prove just *two* wqo theorems to complete the analysis, namely those for which the forbidden subtournament is either a specific tournament of order 5, or another specific tournament of order 6 (the latter tournament can occur in two dual forms, differing only in orientation, but it suffices to treat one form). The required theorems turn out to be correct, and the analysis comes to an end. Each of the two wqo theorems requires a close and relatively lengthy analysis, given separately in [34, 35], and of a wholly different character from all that has gone before. The main tool at this stage is Kruskal's tree theorem and a good deal of direct analysis (the more recent draft of [35] also makes good use of [48]).

We emphasize that before one undertakes the proofs of the structure theorems given in [34, 35], one has considerable confidence that a sufficiently close analysis will either produce the required structure theorems, or reveal a more subtle construction of an additional antichain, and one has a theoretical guarantee that if one continues in this fashion long enough, the process must terminate. For the case we have described we had only finitely many cases to consider, but in general we have a wqo set of problems to handle, and no specific upper bound on how many iterates will be required. But we are assured that the dialectic must come to an end, and we will arrive at utopia, or in any case at a stopping point.

In the case $k = 1$, one has the anticipated structure theorems. In both cases the tournaments excluding the given forbidden subtournament, which has order 5 or 6, can be analyzed as built up along a tree using comprehensible components at each stage, where the pieces involved are considered comprehensible if they come from sets which are obviously wqo under embedding, such as linear orders or tournaments of a fixed bounded size.

In this direction, some of the remarkable work associated with the proof of the Graph Minor Theorem may be relevant; we are looking for tree decompositions of tournaments, and if that side of the picture could be systematized further in our context, then these analyses would flow much more smoothly.

With Λ_1 identified, and with a proof of its correctness in hand, we have the decidability of Problem A for tournaments, in the case of two constraints, something realized after the fact in [9]. Our abstract knowledge of decidability always marches one step ahead of our ability to say anything concrete about the problem, and, in particular, gives us a clear framework for the next step of the analysis.

We took this further in [9], finding three infinite families of isolated antichains (of growing complexity in terms of the sizes of the associated constraint sets), and, in particular, we found a candidate for Λ_2 which is reasonable as a first try—this set contains only three new antichains in addition to those of Λ_1, one of which is Henson's original example. We can say something more about the construction of antichains—the known ones can be viewed as "imported" from a simpler combinatorial setting, which we examine in §2.5.

At this point, if one believes that the candidate for Λ_2 is correct, this means that for any pair of constraints not ruled out by the known antichains, we anticipate a structure theorem leading to a wqo result. This amounts to a cornucopia of conjectured structure theorems, an infinite series, unlike the previous case where there were, a priori, only finitely many instances left to examine after the first candidate for Λ_1 was put forward. The evidence for these specific conjectures is not particularly strong, other than the finiteness theorem itself, which suggests that something qualitatively like this picture should be correct.

2.3. Problem A for graphs

The wqo problem was taken up originally in the context of tournaments for much the same reasons that Henson originally constructed one such antichain: the analysis of homogeneous directed graphs.

A directed graph G is *homogeneous* if any isomorphism $\alpha_0 : A \cong B$ between two of its finite substructures is induced by an automorphism α of G. This is a highly restrictive condition, but using a classical construction of Fraïssé, Henson [21] showed that any antichain of tournaments translates into uncountably many homogeneous digraphs (embedding some of them, and omitting others, at random). Later it turned out that there are not so many other ways to build homogeneous directed graphs, and in fact once the Henson technique has been exploited, there remain only countably many further examples, which can be explicitly identified [8]. As a corollary, a variety of simple questions about these homogeneous directed graphs translate back to the structure of the quasiordered class $\mathcal{Q}$ of finite tournaments, and its finitely constrained ideals. In particular one such question, concerning the number of homogeneous digraphs omitting finitely many specified digraphs, translates directly into the wqo problem considered here, for the class $\mathcal{Q}$.

But no doubt the problem has a similar character over a broad range of combinatorial contexts, and with few exceptions the issue of decidability is probably much the same, independent of the particular context. The finiteness theorem certainly applies.

We will consider two other cases of some interest: the case of graphs, and the case of permutation patterns. The case of graphs was treated in [16], and in the case of permutation patterns there is a nice theory, still not complete, which has made considerable progress recently. The problem degenerates in the case of graphs, and to get something of the expected level of complexity one would need to introduce a little more structure, such as a coloring of the vertices by two colors, or any finite number greater than 1.

In the case of graphs, the first antichain that comes to mind is the collection of cycles $I_0 = \{C_n : n \geq 3\}$. We claim that this single antichain gives us Λ_k for all k. The antichain I_0 is certainly minimal, as any $J \leq I_0$ which is not simply a subset of I_0 would consist, apart from finitely many elements, of subgraphs of paths, and it is easy to see that these cannot form an antichain; to see this, it is helpful to encode these graphs by strings giving the lengths of their connected components. The natural partial order on these strings, or more generally on any strings taken from a quasiordered alphabet, is the following:

Definition 2.3. Let $\mathcal{Q}$ be a quasiorder, $\mathcal{Q}^*$ the set of finite strings with elements in $\mathcal{Q}$. For $s = (s_i)$, $t = (t_j)$ in $\mathcal{Q}^*$, we write $s \leq t$ if there is an increasing function $i \mapsto j_i$ such that

$$s_i \leq t_{j_i}$$

for all i.

In our case, the strings are strings of natural numbers, and an embedding between two such strings—in this sense—gives an embedding of the corresponding graphs. Furthermore by a result of Higman [22], if $\mathcal{Q}$ is wqo then $\mathcal{Q}^*$ is wqo; so with $\mathcal{Q}$ the natural numbers under their usual ordering, our claim follows.

The result of Higman referred to here was mentioned earlier for the case of strings of fixed length, just before the proof of the finiteness theorem. It is a special case of Kruskal's tree theorem, and is equivalent to the case in which the trees involved have height 1 (or, applying that version several times, the case in which they have fixed finite height).

However we claim not merely that I_0 is a minimal antichain, but that it will serve for Λ_k for any k, or in other words: if a graph C embeds in infinitely many cycles, and $\mathcal{C}$ is a finite set of constraints containing C, then the class of graphs $\mathcal{Q}_{\mathcal{C}}$ is wqo.

Evidently, the graph C above may be supposed to be a path. The structural analysis of graphs omitting a path of fixed length was given in [28] with an eye toward proving the existence of the corresponding universal graph, and in [16] with a view toward the wqo problem. The proof involves strengthening the claim a bit and then proceeding inductively; and the strengthened form of the claim turns out to have additional uses, so we will present this in detail.

We consider graphs with a vertex coloring c in a wqo alphabet Σ (in other words, an arbitrary function from the set of vertices to Σ). In this context, an embedding between two colored graphs (G, c) and (G', c') would be an ordinary embedding of $f : G \to G'$ as a subgraph, respecting the coloring in the sense that $c'(f(v)) \geq c(v)$ for v in G. In the frequently occurring case in which Σ is finite and all elements of Σ are incomparable, this condition reduces to $c'(f(v)) = c(v)$.

Proposition 2.4 ([16])**.** *Let P be a fixed finite path and Σ a wqo alphabet. The class of graphs equipped with a vertex coloring by Σ which omit the path P is wqo under embedding. In particular, the class of graphs which omit P is wqo under embedding as a subgraph.*

Proof. Proceed by induction on the length n of P. It suffices to deal with structures in which the underlying graph G is connected.

By induction, we may set aside those vertex colored graphs in which there is no path of length $n-1$. So we consider the structure of a connected finite colored graph G, not containing a path of length n, but containing some path P_0 of length $n-1$. We fix one such path P_0 together with an enumeration of its vertices. We break $G \setminus P_0$ into its connected components, and pass to a larger color set in which every vertex receives a pair of colors: first, its original color in G; and second, the set of vertices in P_0 to which it is adjacent, coded as a subset of $\{1, \ldots, n-1\}$. Thus the enhanced coloring uses an alphabet of the form $\Sigma \times \Sigma'$ with Σ' finite.

Now any connected component of $G \setminus P_0$ does not embed a path of length $n-1$, as otherwise we have two disjoint paths of length $n-1$ in G, and then a path of length n, by inspection. So by induction, the connected components of $G \setminus P_0$ come from a wqo set under embedding; call this wqo set Σ_{n-1}, and view it as an alphabet. Then G itself can be encoded by a finite string with entries in Σ_{n-1}; here the order of the terms is unimportant, but we may fix an order, and from the string we can uniquely reconstruct G. And indeed from an embedding of one such string into another, in the sense just described above, we get an embedding between the corresponding colored graphs, and so by Higman's theorem [22] we again have a wqo set. □

Corollary 2.5. *We may take Λ (i.e., $\bigcup_k \Lambda_k$) to be $\{I_0\}$; that is, if a finite set of constraints allows an antichain, then it allows a cofinite subset of I_0.*

Proof. Otherwise, one of our constraints embeds into a path, and hence the corresponding ideal $\mathcal{Q}_C$ is wqo by Proposition 2.4. □

In spite of Corollary 2.5, there is another minimal antichain of graphs, consisting of a set of trees. The so-called *arrows* or *bridges,* are trees consisting of two vertices of degree 3, joined by a path. Let this antichain be called I_1. Evidently I_1 is a minimal antichain: proper subgraphs of these trees have as their connected components paths or paths extended by one vertex of order 3. So again the components form a wqo set, and by Higman's theorem the corresponding graphs do as well.

As this antichain is not actually needed to form Λ, it gives us a simple example of a non-isolated antichain (in a topological space with only two points). We have not yet encountered such examples in the context of tournaments.

Ding shows that these two antichains characterize the downward closed sets of graphs which are wqo; in other words, up to a natural notion of equivalence, these are the only minimal antichains.

First, we clarify the notion of equivalence between minimal antichains defined above by:

$$I \sim I' \text{ if and only if } \mathcal{Q}^{<I} = \mathcal{Q}^{<I'}$$

Lemma 2.6. *Let I, J be minimal antichains. Then $I \sim J$ if and only if $I \cup J$ is a minimal antichain (identifying points q of I with any equivalent points q' in J, that is points with $q \leq q' \leq q$). In particular, if there is an infinite antichain A with $A \leq I, J$, then $I \sim J$.*

Proof. Suppose $I \sim J$. Then no $q \in I$ belongs to $\mathcal{Q}^{<J}$, and vice versa, so $I \cup J$ is an antichain. Suppose $A \leq I \cup J$ is another infinite antichain. Let $A_I = \{a \in A : \exists q \in I \; a \leq q\}$ and define A_J similarly. We may suppose that A_I is infinite. Then $A_I \subseteq I$, by minimality. Now consider $a \in A_J$, and $q \in J$ with $a \leq q$. Suppose $q \not\leq a$. Then $a \in \mathcal{Q}^{<J} = \mathcal{Q}^{<I}$. So $\{a\} \cup A_I \leq I$ and thus $a \in I$ up to equivalence. Thus $A_I \subseteq I$, $A_J \subseteq I \cup J$, and we conclude.

Now suppose that $I \cup J$ is a minimal antichain (after making appropriate identifications) and take $a \in \mathcal{Q}^{<I}$. Let $A = \{q \in I \cup J : a \not\leq q\}$. If A is infinite, then $A \cup \{a\}$ is an antichain and $A \cup \{a\} \leq I \cup J$ forces $a \in I \cup J$. If $a \in I$ then I is not an antichain. If $a \in J$ and $a \leq q' \in I$ then as $I \cup J$ is an antichain, a and q' must be equivalent. Thus we return to the case $a \in I$ to get a contradiction.

So A is finite and in particular there are at least two $q, q' \in J$ so that $a \leq q, q'$. Hence $a < q, q'$ and $a \in \mathcal{Q}^{<J}$.

For the final point, if $A \leq I, J$ is infinite, then $A \subseteq I \cap J$ and $I \sim A \sim J$. □

Fact 2.7 ([16, Theorem 2.7])**.** *The only minimal antichains for the case of graphs are I_0 and I_1, up to equivalence.*

In other words, the claim is that if a downward closed class of graphs contains no large cycles and no large arrows, then it is wqo (even, in fact, with respect to the induced subgraph relation).

Ding proceeds roughly as follows. Let G be a connected graph which is not itself a path. For any vertex v of G of degree at least 3, remove from G all components of $G \setminus \{v\}$ which are paths, and label v by a sequence of natural numbers consisting of the orders of all the removed paths, in some order. The result is a labeled graph G' whose labels are finite sequences of natural numbers. If G is a path, then let G' be a single vertex labeled by the length of that path.

Then for any two connected graphs G, H, an embedding of the labeled graph G' into H' gives rise to an embedding of G into H. So it suffices to show that for fixed N, the labeled graphs G' associated to connected graphs G omitting all cycles and arrows of order at least N form a wqo set; and for this, Proposition 2.4 suffices. In the first place, the set of labels is a wqo set. In the second place, it turns out that the reduced graphs G' contain no paths of length $3N$, as one can see by considering a path P of maximal length in G', and attempts to extend P further at one end or the other.

Wqo problems relative to the partial ordering of embeddability as an *induced subgraph* were also considered in [16, 42]; here the constraints are forbidden induced subgraphs. We are not aware of any very systematic attempt to identify the minimal antichains relevant to finite sets of constraints, in this context. On the other hand, a considerable strengthening of wqo was considered by Pouzet in [43]: he considers downward closed collections of graphs which are wqo, and remain wqo if one allows arbitrary vertex colorings by n colors. Call such classes *n-wqo*. Pouzet conjectures that 2-wqo classes are already n-wqo for all n. An easy argument shows that all such classes are determined by finitely many constraints, which is not true of wqo classes in general. This is discussed in detail in [15], along with a systematic approach to Pouzet's conjecture.

2.4. Problem A for permutations

We deal here with (finite) permutations omitting certain patterns. As Cameron observes in [7], this falls under our structural point of view by considering a permutation to be encoded by a pair of linear orderings; the isomorphism types of permutations are then *permutation patterns.* The study of such permutations arises naturally in the theory of sorting; in particular, permutations that can be sorted back to standard order using a stack (last-in, first-out) are those omitting the pattern (231), and the number of such permutations on a set of order n is given by the Catalan numbers [25].

The theory has grown considerably, with considerable emphasis on enumeration (explicitly, or asymptotically), as well as the connection with computational issues (such as more elaborate sorting devices). See [3] for a survey. For some naturally occurring downward-closed classes (such as those associated with particular sorting mechanisms) it is not immediately clear that there is a characterization by a finite set of forbidden patterns: in other words, the minimal unsortable permutations could possibly form an infinite antichain. So the study of infinite antichains of permutations naturally accompanies the subject, and is the subject of Chapter 7 of [3].

One can find a discussion of recent work on the structure of minimal antichains of permutations in the thesis of Waton [52] and a survey talk by Brignall [5], and we will go over some of this together with a little ancient history. Our Decision Problem $A_{\mathcal{Q}}$, for permutations, is given as *Algorithmic Problem 2.7.5* of [52]. Also worthy of note here is the Enumeration Problem, given as Algorithmic Problem 2.6.4: if a downward closed class has a finite description, is its enumeration function computable in polynomial time? (In [52] the notion of finite description is understood broadly; it certainly includes any specification by finitely many forbidden patterns.)

The first antichain given in [52, p. 35], omits decreasing sequences of length 3 (i.e., the pattern (321)). It can be derived from the zigzag, which for odd length n would be:

$$\sigma_n = (3, 1; 5, 2; 7, 4; 9, 6; \ldots; n, (n-3); (n-1))$$

We may replace the initial pair $3, 1$ and the final pair $(n-3), (n-1)$ by a pattern of type 2341, relabeling the remaining terms to avoid clashes, as follows

$$\sigma'_n = (2351; 7, 4; 9, 6; 11, 8; \ldots; (n+2); (n-1), (n+3), (n+4)(n+1))$$

Under an embedding of one such permutation σ'_n into another σ'_m, the index 1 goes to an index preceded by three larger ones; so it goes to itself, and the map is the identity on the "anchor" consisting of the first four entries, and once the pair $5, 1$ are fixed, then so are 4 and 7, 6 and 9, etc., till at the end a contradiction is reached if $m \neq n$.

Evidently, the structure of this antichain is that of an oriented path with the ends colored.

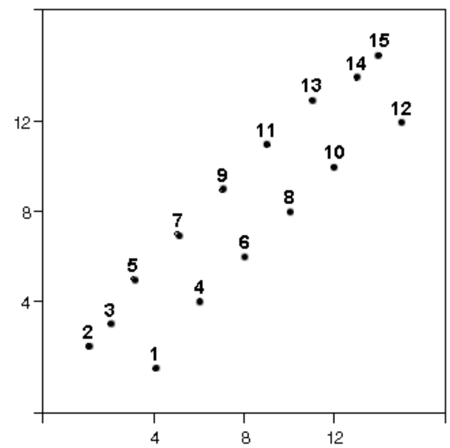

Figure 1: σ'_{11}

Just as in the case of tournaments, the set Λ_1 (the finite set of minimal antichains needed to settle all wqo problems for the case of one constraints) has been fully identified for the quasiordering on permutations, and consists of three antichains, of which the one shown is the most straightforward [1], leading to the result that for a single constraint α, the set of permutations avoiding the pattern α is wqo if and only if α is one of the following:

$$(1), (12), (21), (132), (213), (231), (312)$$

It also follows that the problem *to decide whether a closed class of permutations determined by the exclusion of two given patterns is wqo* is decidable in polynomial time. But as far as I know there is as yet no known algorithm, and for that matter no bound on the degree of the relevant polynomial.

There is also a very elegant and general construction of a variety of minimal antichains (known as "fundamental antichains" in this neck of the woods) in a systematic way, more subtle than the method we will describe in §2.6. This is based on the two-dimensional nature of permutations when one encodes them as a pair of linear orderings (equivalently, in terms of the graph of σ as a subset of the plane, with the axes ordered).

The permutation is then a scattered set of points lying within a square. If one imposes a grid structure on this square with a fixed number of boxes, and requires that the permutation misses some boxes, and meets the remainder in monotonically increasing or decreasing sequences, then with the grid fixed we get a large number of downward closed classes of permutations. Furthermore the grid structure is encoded by a matrix with entries ± 1 where the permutation is required to be monotonic, and 0 where it is required to be absent.

Murphy and Vatter give an explicit criterion for such a matrix to define a wqo set of permutations in terms of a graph derived from the matrix: the associated class is wqo if and only if the graph contains no cycles. Furthermore, the proof of the failure of wqo is by an explicit construction of minimal antichains which "wind around" such a cycle [39, 52, 5].

To quote one more point from [52], while the wqo problem for classes of permutations defined by grid constraints has been solved, whether these classes are themselves defined by finitely many constraints is open. Quoting from §4.9: "The basis problem is particularly frustrating. It is very natural to conjecture that every grid class is finitely based, see for example Huczynska and Vatter [24, Conjecture 2.3]. ... Nonetheless, a proof is not only elusive, even an approach that hints at the beginnings of a proof has not been found."

As a point of history, I quote a comment by the authors of [49] (linked to their paper online) that the earliest known examples of infinite antichains of permutations appear to be constructions by Tarjan, Pratt, and Laver in 1972, 1973, and 1976 respectively [50, 44, 36].

As Cameron also pointed out in [7], once one sees permutations as structures equipped with two linear orders, it is natural to take this as a model for the study of more complex structures equipped with k linear orders, $k \geq 2$ fixed. And this line is taken up as well in [52] (§5.9).

2.5. Problem A for colored paths

In the present subsection we will consider the quasiorder $\mathcal{Q}^{(c)}$ whose elements are finite oriented paths equipped with a coloring of the vertices using colors taken from the set $\{1, \ldots, c\}$ with c fixed. We aim to show that the isolated, minimal antichains are dense, and each isolated antichain is effective (that is, membership in the corresponding ideal $\mathcal{Q}^{c,<I}$ is algorithmically decidable). And we claim that for $c = 2$, the class $\mathcal{Q}^{(2)}$ embeds into the quasiorder of finite tournaments in such a way that its minimal antichains are carried to isolated, minimal, and effective antichains of the class of finite tournaments.

We encode the elements of $\mathcal{Q}^{(c)}$ by words in the language $\{1, \ldots, c\}^*$ (arbitrary words in the alphabet $\{1, \ldots, c\}$). In this language, the embeddings to be considered are embeddings of one word as a contiguous segment of another.

Definition 2.8. Let a, a' be words in the alphabet $\{1, \ldots, c\}$, of length k, with a' a cyclic permutation of a.

1. A word $w \in \{1, \ldots, c\}^*$ is (a, a')-*periodic* if w begins with a, ends with a', and is k-periodic.

2. For $i, j \in \{1, \ldots, c\}$ let $I_{ij}(a, a')$ be

$$\{(i) \frown w \frown (j) : w \text{ is } (a, a')\text{-periodic}\}$$

   We will write these elements more briefly as iwj.

3. The pair $i, j \in \{1, \ldots, c\}$ is *appropriate* to the pair a, a' as above if ia and $a'j$ are *not* k-periodic, that is $i \neq a_k$, and $j \neq a'_1$.

Lemma 2.9. *For a a word of length k in $\{1, \ldots, c\}^*$, a' a cyclic permutation of a, and i, j appropriate to the pair a, a', the set $I = I_{ij}(a, a')$ is an isolated minimal antichain in $\mathcal{Q}^{(c)}$, and the corresponding ideal of $\mathcal{Q}^{(c)}$ is effective.*

Proof. First, I is an antichain. An embedding of iwj into $i \frown w' \frown j$ will send iw into iw', and as iw is not k-periodic it will carry i to the first entry and thus iw goes over to an initial segment. Similarly wj goes into $w'j$ as a terminal segment and thus $w = w'$.

Now minimality holds since the sequences (iw) and (wj) associated with I are increasing.

The fact that these antichains are isolated is an expression of their almost periodic structure; with finitely many forbidden sequences, one can pin down that structure.

And the effectivity is clear on similar grounds. □

We will call these particular antichains "almost periodic," of period k.

While we do not claim to have a complete understanding of the minimal antichains in $\mathcal{Q}^{(c)}$, the following gives us everything we need.

Proposition 2.10. *Let $C \subseteq \mathcal{Q}^{(c)}$ be finite. If $\mathcal{Q}^{(c)}_C$ is not wqo, then it contains an almost periodic antichain.*

Proof. Let $I \subseteq \mathcal{Q}_C^{(c)}$ be an infinite antichain. Let $k_0 = \sup(|P| : P \in C)$. Set $K = k_0(c^{k_0}+1)$. For $A \in I$, let A^L be the longest initial segment of A which is k-periodic for some $k \le K$, and let A^R be the longest terminal segment of A disjoint from A^L which is k-periodic for some $k \le K$.

Thinning I, we may suppose that the terms A^L for $A \in I$ are increasing, with each occurring as a terminal segment of the next, and that similarly the terms A^R increase, with each an initial segment of the next. With A^L on the left and A^R on the right under control, we consider the middle part $\hat{A}$: $A = A^L \hat{A} A^R$. These middle terms are all distinct since I is an antichain, and in particular their length is unbounded.

Consider $A \in I$ for which $|\hat{A}| \ge K$. Considering the first $c^{k_0}+1$ successive disjoint paths in $\hat{A}$ of length k_0, we find two disjoint occurrences in $\hat{A}$ of the same path of length k_0. In the notation of words, $\hat{A}$ contains a segment $ww'w$ with $|w| = k_0$ and with w' possibly empty. Let $a = ww'$. Then a^n is C-free for any n since $ww'w$ is. In other words, for any cyclic permutation a' of a, any (a, a')-periodic word is C-free.

We claim now that there are cyclic permutations a', a'' of a, and elements $i, j \in \{1, \ldots, c\}$, so that ia' and $a''j$ are C-free and are not k-periodic. Suppose the contrary, and specifically that this fails on the left: for any cyclic permutation a' of a, and any $i \in \{1, \ldots, c\}$ other than the final term a'_k of a', the word ia' is not C-free. Then this forces the initial segment of A up through any occurrence of a in $\hat{A}$ to be k-periodic, and contradicts the choice of A^L.

So with a', a'' and i, j as above, the almost periodic antichain $I_{ij}(a', a'')$ lies in $\mathcal{Q}_C^{(c)}$. □

In the statement of the next corollary we make use of the topological language touched on earlier. In particular, a set $\mathcal{I}$ of minimal antichains is called *dense* if any ideal in $\mathcal{Q}$ defined by finitely many constraints which is not wqo contains an antichain equivalent to one in $\mathcal{I}$. That is, the finite sets of constraints define the basic open sets, and the nonempty basic open sets meet $\mathcal{I}$, up to equivalence.

Corollary 2.11. *The isolated minimal antichains for $\mathcal{Q}^{(c)}$ are exactly the almost periodic antichains. They are dense and their ideals are uniformly effective in the sense that the relation "$x \in \mathcal{Q}^{<I_{ij}(a,a')}$" is decidable as a relation in x, i, j, a, a'. The determination of the finite constraint sets C for which $\mathcal{Q}^{(c)}$ is wqo is effective (algorithmically decidable). The correspond-*

ing set $\Lambda^c = \bigcup_n \Lambda^c_n$ for $\mathcal{Q}^{(c)}$ can be taken to be the set of almost periodic antichains, and this is the minimal choice possible.

Proof. The previous proposition gives density, and we know these antichains are isolated. Their structure is so simple as to make the uniform effectivity clear. For the decidability of the wqo problem one must determine effectively whether a given constraint set allows an almost periodic antichain. By our proof, if there is a $\mathcal{C}$-free almost periodic antichain, then there is one whose period is at most $K = k_0(c^{k_0} + 1)$. So the problem is a finite one.

The last assertion holds (for the set of isolated minimal antichains) whenever the isolated minimal antichains are dense. □

2.6. From colored paths to tournaments

The classes $\mathcal{Q}^{(c)}$ provide more than a convenient case study: they are readily encoded into other contexts, and give our "standard model" for the construction of minimal antichains. It remains to make this last point explicit. Before turning to concrete examples, let us consider what sort of encoding is wanted.

A natural way to embed $\mathcal{Q}^{(c)}$ in the quasiorder of tournaments is as follows. First find a sequence of tournaments T_n for n varying through an infinite index set X, so that T_n has vertex set $\{1, \dots, n\}$, and so that any embeddings $T_m \to T_n$ for $m, n \in X$ must be a shift map $x \mapsto x + k$ from $\{1, \dots, m\}$ to $\{1, \dots, n\}$, as in the case of oriented paths. Represent the colors in $\mathcal{Q}^{(c)}$ by binary strings of length k where $2^k \geq c$. For each vertex colored path P_n on $\{1, \dots, n\}$, let the corresponding tournament $T(P_n)$ be obtained from T_n by adjoining k vertices $v_1, \dots, v_k$ with some fixed structure, e.g. a linear ordering with $v_1 \leq v_2 \leq \cdots \leq v_k$, and using the coloring of $\{1, \dots, n\}$ to determine the edge relations between the vertices v_i and the vertices of T_n. In other words, if $1 \leq i \leq k$, and $1 \leq j \leq n$, let $v_i \to j$ if and only if the color $c(j)$ associated to j corresponds to a bit string s for which $s(i) = 1$. In particular, for $c = 2$, one additional vertex suffices.

Example 1. *Let L_n be the natural linear order on $\{1, \dots, n\}$, viewed as a tournament, and let P_n be the result of reversing the edges $(i, i+1)$ in L_n. Included in Fact 2.13 below is the claim that for $n, n' \geq 6$, all embeddings from P_n into $T_{n'}$ are translation maps from $\{1, \dots, n\}$ to $\{1, \dots, n'\}$. If P is an oriented path of order n with a vertex coloring by 2 colors, denoted $+$ and $-$, then $T(P)$ denotes the extension of T_n by a single vertex v_0 in* *which the orientation of the edges from v_0 to the vertices of T_n is governed by the coding. Here we identify the vertices of P or T_n with $\{1,\ldots,n\}$, and in particular with each other.*

Returning to the general case, for m, n large and for most vertex colorings, one expects that an embedding of $T(P_m)$ into $T(P_n)$ will send the set $\{1,\ldots,m\}$ into $\{1,\ldots,n\}$, in which case it will also send each of the added vertices v_i to itself if the structure on these vertices is rigid (as in the case of a linear order on the v_i). In such cases we will get an embedding of $T(P_m)$ into $T(P_n)$ if and only if there is such an embedding $P_m \to P_n$, respecting the coloring. Thus if (P_n) is an antichain in $\mathcal{Q}^{(c)}$ one expects, after dropping a few terms, that $(T(P_n))$ will be an antichain. If (P_n) is almost periodic and if the set (T_n) is itself isolated by finitely many conditions then the $T(P_n)$ will be isolated, and have whatever effectivity properties the sequence (T_n) has. There remains the question of the transfer of minimality: if (P_n) is a minimal antichain in $\mathcal{Q}^{(c)}$, is $(T(P_n))$ also minimal?

We need to consider the effect of removing one of the vertices v_i from each of the tournaments $T(P_n)$. In terms of P_n, this involves a collapse of the color set, in which certain pairs of colors become identified. For $k = 1$, $c = 2$ this is not an issue since removal of v_1 leaves us with T_n in that case. For $k > 1$ it is an issue. We explore this further.

Suppose we begin with an almost periodic antichain $I_{ij}(a, a')$ in $\mathcal{Q}^{(c)}$, and we use an encoding procedure with k auxiliary vertices to convert this into an antichain of tournaments. Here a and a' have length ℓ, and we have the conditions $i \neq a_\ell$, and $j \neq a'_1$. If we identify some colors, but avoid identifying i with a_ℓ or j with a'_1, then we again have an antichain, involving fewer colors. If removal of some auxiliary vertex i corresponds to such an identification of colors, then our antichain encoding $I_{ij}(a, a')$ is not minimal, and contains a minimal antichain encoding an antichain $I_{\bar{\imath},\bar{\jmath}}(\bar{a}, \bar{a}')$ involving fewer colors.

In particular, if the colors i and a_ℓ are encoded by strings of length k differing in at least two places, removal of an auxiliary vertex v_i will not identify them. If $k \geq 3$ it follows that there is some vertex v_i which can be removed without collapsing either pair of colors (i, a_ℓ) or (j, a'_1). So in our encodings of paths by tournaments we may take $k = 2$ and correspondingly $c \leq 4$. Each of the corresponding antichains is either minimal, or lies above a minimal antichain corresponding to an encoding with $k = 1$ and $c = 2$.

Now we make this more concrete, and we deal first with the construction of appropriate tournaments (T_n).

Notation 2.12.

1. *$L_{k,n}$ is the tournament with vertex set $\{0, \ldots, n-1\}$, and with edges determined by this rule: for $i < j$, the pair (i, j) is an arc if and only if $j \equiv i \mod k$.*

2. *$N_{k,n}$ is the tournament obtained from $L_{k,n}$ by reversing the orientation of each arc connecting successive vertices $(i, i+1)$.*

We should explain the idea. Begin with a linear order whose vertices are colored by k colors; specifically, let the vertex set be $\{0, \ldots, n-1\}$ and take the residues mod k as the colors. Encode this structure by a tournament as follows: within each class, leave the edge relation alone; between distinct classes, reverse it. This gives $L_{k,n}$.

After that, $N_{k,n}$ is derived from $L_{k,n}$ by reversing precisely those edges which correspond to the successor relation in the original structure. This is an attempt to make the successor relation more "visible," that is, more likely to be preserved under embeddings from one of these tournaments to another.

Fact 2.13 ([9])**.** *Embeddings from $N_{k,n}$ to $N_{k',n'}$ are translation maps (with $k = k'$) in the following cases:*

1. *$k = 1$ and $n \geq \max(6, 2k' + 1)$;*

2. *$k = 2$, $n \geq 6$;*

3. *$k \geq 3$, and $n \geq 6k + 1$.*

This gives us an ample supply of tournaments $T_n = N_{k,n}$ for our purposes. The following seems quite likely, but we have not looked into it at this level of generality. Some special cases were given in [9], but at that point we had not looked separately into $\mathcal{Q}^{(c)}$.

Conjecture 1. *Let I be an almost periodic antichain in $\mathcal{Q}^{(2)}$, and k fixed. For $P \in I$ let $T_k(I)$ be the corresponding tournament using the sequence $(N_{k,n})$ as the base. Then after removal of finitely many terms, $(T_k(P) : P \in I)$ is an isolated minimal antichain in the quasiorder of finite tournaments, whose associated ideal is effective.*

The minimality is not at issue and the isolation and effectivity do not seem problematic, though there is something to work out. But the main thing to check is that whatever unnatural (sporadic) embeddings there may be between $T_k(P)$ and $T_k(P')$ for small P, P', eventually die out.

Most of the antichains of tournaments given to date fall into this category. In [9, Proposition 5.7] three antichains built as 1-point extensions of N_n (i.e., $k = 1$) were given, corresponding to the antichains $I_{ij}(a, a')$ of the following forms:

1. $a = 0^m 1^m$, $i = 0$; $a' = 0^m 1^m$, $j = 1$ or $a' = 1^m 0^m$, $j = 0$.
2. $a = 01^m 0^{m-1}$, $i = 1$; $a' = 0^{m-1} 1^m 0$, $j = 1$ or $a' = 1^{m-1} 0^m 1$, $j = 0$.
3. $a = 1$, $i = j = 0$.

In [9, Proposition 5.3] two specific families of isolated minimal antichains were given, built from $N_{k,n}$ with k arbitrary. One of these also falls into our current framework.

Notation 2.14. *Let $N_{k,n,H}$ be the variant of $N_{k,n}$ in which the orientation of the arc connecting the extreme points 0 and $n - 1$ is reversed.*

For $k = 1$ this is the construction given by Henson [21], and in [9] it is shown (sketchily) that $(N_{k,kn+1,H} : n \geq 6)$ is an isolated minimal antichain.

As only the arc connecting $0, n - 1$ is reversed, this can be viewed as a 1-point extension of $N_{k,kn}$ by the point kn. The periodic words involved are $a = 0^{k-1}1$ and $a' = 010^{k-2}$ if $k \geq 2$, and just $a = a' = 1$ otherwise (as in the third case of Proposition 5.7).

This leaves one more antichain from [9] to be accounted for. This one comes from an even more direct encoding.

Notation 2.15.

1. *If A is a tournament and v a vertex of A, then the tournament A^v obtained by* doubling *the vertex v has one additional vertex v^*, and for $u \in A$, we take $u \to v^*$ iff $u \to v$.*
2. *The tournament $N_{k,n,D}$ is obtained from $N_{k,n}$ by first doubling 0, then doubling $n - 1$.*

Again, $(N_{k,kn+1,D} : n \geq 6)$ is an isolated minimal antichain. This does not fit into our framework of encoding $\mathcal{Q}^{(c)}$. Really what we are encoding are paths P_n with the ends marked by constants u_0, u_1, a simpler sort of antichain. Indeed, there is a natural congruence on $N_{k,kn+1,D}$ defined by: $u \sim u'$ if for all $v \neq u, u'$ we have $u \to v \iff u' \to v$. There are two classes of order 2, the remainder of order 1, and the quotient is isomorphic to $N_{k,kn+1}$. So in a weak sense the endpoints are "marked" by being doubled. That this actually gives an antichain does not immediately follow by general principles, so one uses the embedding properties of the $N_{k,n}$ to check it. However, it is clear from the use of the doubling construction that if it is an antichain, it is minimal. One could presumably repeat this, given other tournaments with the properties of the $N_{k,n}$.

We recall the following.

Fact 2.16 ([33]). *Λ_1 may be taken to consist of $I_1 = (N_{1,n,D} : n \geq 7)$ and $I_2 = (N_{2,2n+1,H} : n \geq 4)$.*

These two antichains originally appeared as modified orders and modified local orders, respectively, in other words they are derived from linear orders, and linear orders with a coloring of the vertex by two colors, respectively. To date, all known minimal antichains of finite tournaments are modest generalizations of these two, as described above.

The fundamental conjecture for those in an optimistic frame of mind, would be the following.

Conjecture 2. *Within the quasi-order of finite tournaments with respect to embeddings, the isolated minimal antichains are dense, and the associated ideals are uniformly effective. The determination of whether a given finite set of constraints is compatible with an isolated minimal antichain is also decidable, so Problem $A_{\mathcal{Q}}$ is decidable, for tournaments.*

We see nothing unreasonable in this. One may of course read "permutation" in place of "tournament" here and get a conjecture which appears to have much the same force.

But in the case of tournaments, we have noticed that the known facts are compatible with the stronger statement that all of the minimal antichains come from natural encodings of known antichains in simpler classes $\mathcal{Q}$. We consider the notion of isolation as the key here, though it may in practice work out to some form of almost periodicity in this particular context. One can perhaps read the "grid" theory of permutation classes as also involving a coding of colored paths by permutations (via "pin sequences" and symmetry operations [5]).

An interesting question is whether we can find a direct encoding of the known permutation antichains back into the quasiorder of tournaments. This could give examples of (isolated) minimal antichains of tournaments quite different from any previously encountered.

In the current state of knowledge, one may freely conjecture similar things for any natural class of finite combinatorial structures. But if one generalizes sufficiently far, using the methods of computability theory, one encounters extreme examples of undecidability (cf. §4.1), which may or may not become relevant once one deals with very rich combinatorial structures. And for that matter, there is nothing in the current state of knowledge to prevent such phenomena from arising in either of the cases of tournaments or permutation patterns. Still, we think this last possibility is highly unlikely, and we'll place our current bets on Conjecture 2.

3. The universality problem with constraints

The subject of the present section is Problem B for graphs: given a finite collection $\mathcal{C}$ of finite connected graphs, determine whether there is a universal $\mathcal{C}$-free graph.

As the constraints in $\mathcal{C}$ are taken to be connected, a disjoint sum of $\mathcal{C}$-free graphs is $\mathcal{C}$-free. Hence, if we have a countable family (G_i) of jointly universal countable $\mathcal{C}$-free graphs—meaning, that any $\mathcal{C}$-free graph embeds into one of these as an induced subgraph—then we also have a single universal countable $\mathcal{C}$-free graph, their direct sum. So to prove non-universality one would look for a construction of uncountably many pairwise incompatible countable $\mathcal{C}$-free graphs. This approach is not only natural, but inevitable, as we shall see.

3.1. Graphs of bounded degree and other special cases

Among the classes of graphs determined by finitely many forbidden substructures, those in which the graphs have bounded vertex degree (that is, where a star is included among the constraints) can be analyzed in a straightforward manner. In fact we can show that the cases in which a *weakly universal* graph exists are severely limited. This is one of the exceptional situations in which weakly universal graphs are more common than strongly universal graphs. Experience suggests that typically, when there is a weakly universal graph then there is also a strongly universal one (though possibly less obviously: e.g., an infinite complete graph is weakly universal for the class of all countable graphs, while a strongly universal one actually requires some construction).

In the case of bounded degree, one focuses on the *maximal connected* $\mathcal{C}$-free graphs; these are graphs such that any embedding into a connected $\mathcal{C}$-free graph is an isomorphism. It is not hard to see that any connected $\mathcal{C}$-free graph extends to a maximal one in this case. Furthermore, if there are only countably many isomorphism types of maximal connected $\mathcal{C}$-free graphs then a universal $\mathcal{C}$-free graph may be formed by taking the disjoint sum of countably many copies of each, while if there are uncountably many maximal connected $\mathcal{C}$-free graphs, then there is no universal countable $\mathcal{C}$-free graph. However, the split between the cases in which there are or are not universal countable $\mathcal{C}$-free graphs is generally much sharper than this. At one extreme, we have the possibility that the connected components of $\mathcal{C}$-free graphs are finite. In that case it is clear that there are only countably many maximal connected $\mathcal{C}$-free graphs, and indeed the maximality is not even needed here. On the other hand, if there is an infinite connected $\mathcal{C}$-free graph, we might expect it to be possible to vary its structure in uncountably many ways, and thus we should generally fall into the second class. One obvious exception to this rule would be the case of graphs of vertex degree at most 2, where there are two isomorphism types of infinite connected graphs, and just one of them is maximal. More generally, we may construct a graph by taking an infinite path and attaching to each vertex a disjoint copy of some fixed finite connected graph, and we may then find a finite set of constraints for which this graph is the unique maximal infinite connected graph. Or varying further, instead of taking a single graph repeated along a path, we may take a finite sequence of such graphs, repeated along a path. All of these examples have the special property that in a connected infinite $\mathcal{C}$-free graph there is a unique infinite path. However this does not yet exhaust the possibilities. So we will now take this case up more systematically, from the beginning.

Let $\mathcal{C}$ be a finite set of connected finite graphs, including some star (a tree consisting of one vertex and $d+1$ adjacent leaves). Thus the $\mathcal{C}$-free graphs have vertex degree bounded by d. A $\mathcal{C}$-free connected graph G is *maximal* if any embedding of G into a $\mathcal{C}$-free connected graph (as a subgraph) is an isomorphism. Note that maximality refers both to the vertex set and the edge set. As mentioned, any connected $\mathcal{C}$-free graph embeds as a subgraph into a maximal connected $\mathcal{C}$-free graph (e.g., by Zorn's lemma, since these graphs are countable or finite). The following is our point of departure.

Lemma 3.1. *Let $\mathcal{C}$ be a finite set of finite connected graphs, including a star. Then the following are equivalent:*

1. *There is a weakly universal $\mathcal{C}$-free graph.*

2. *There are finitely or countably many maximal connected $\mathcal{C}$-free graphs.*

Proof. Clear, on the basis of the foregoing remarks. ☐

As far as the "finite" alternative is concerned, Sam Buss has observed the following.

Lemma 3.2 (Buss)**.** *If $\mathcal{C}$ is a finite set of finite connected graphs, then there is an infinite connected $\mathcal{C}$-free graph of bounded vertex degree if and only if $\mathcal{C}$ contains no path.*

Proof. If $\mathcal{C}$ contains a path, then the diameter of a connected $\mathcal{C}$-free graph is bounded, and hence those of bounded vertex degree are finite.

If $\mathcal{C}$ contains no path, then an infinite path is $\mathcal{C}$-free. ☐

At the opposite extreme, we have the following.

Proposition 3.3. *Let $\mathcal{C}$ be a finite set of connected finite graphs, including some $(d+1)$-star, but no path. If there is a weakly universal countable $\mathcal{C}$-free graph, then $\mathcal{C}$ contains some tree S with at most one vertex of degree 3, and no vertex of greater degree.*

Proof. We suppose $\mathcal{C}$ contains no constraint of the specified type. Roughly speaking, we will vary the lengths of cycles embedding in these graphs. More precisely, we will vary the structure of maximal connected $\mathcal{C}$-free graphs viewed as metric spaces, using cycles for this purpose.

Call a $\mathcal{C}$-free connected graph G *vertex-maximal* if for any connected $\mathcal{C}$-free graph containing G, the vertex sets are the same. Since vertex degrees are bounded, any connected $\mathcal{C}$-free graph G_0 can be extended to a vertex-maximal connected $\mathcal{C}$-free graph by attaching trees to some of its vertices.

Now consider how the graph metric on G changes when a vertex-maximal connected $\mathcal{C}$-free graph G is embedded into a maximal connected $\mathcal{C}$-free graph G^*. Let K be the maximum diameter of a graph in $\mathcal{C}$. For any edge (u, v)

occurring in G^* but not in G, there is a constraint C in $\mathcal{C}$ which prevents us from adjoining an edge at u with a new vertex v^*; so a subgraph C_0 of C embeds into G over u in such a way as to prevent this. Therefore the vertex v must lie on the image of C_0 in G, and hence within distance K of u in the graph metric on G.

It follows that the embedding of a vertex-maximal connected $\mathcal{C}$-free graph into a maximal $\mathcal{C}$-free graph perturbs the graph metric at most by a multiplicative factor of K.

Under our hypothesis on $\mathcal{C}$, we claim that for any set X of natural numbers we can find a vertex-maximal connected $\mathcal{C}$-free graph G such that the nontrivial blocks of G are cycles of diameter K^{2n} for $n \in X$. Then embedding each of these into a maximal connected $\mathcal{C}$-free graph, we can recover X from the metric structure by looking at the metric space analog of cycles.

The construction begins by letting G_X be the disjoint union of cycles of appropriate diameter, joined by paths of length greater than K. Then G_X is $\mathcal{C}$-free in view of our hypothesis on $\mathcal{C}$.

We then extend G_X to a vertex-maximal connected $\mathcal{C}$-free graph G^*_X by attaching some trees to it. After that we pass to a maximal connected $\mathcal{C}$-free graph containing G^*_X, and then varying X we get an uncountable number of nonisomorphic maximal connected $\mathcal{C}$-free graphs. $\square$

In view of Lemma 3.2 and Proposition 3.3, we are left with the case of a constraint set $\mathcal{C}$ containing no path, but containing some tree S with a unique vertex of degree 3, and with no vertex of degree greater than 3; in other words, S is topologically a star whose unique branch vertex has degree 3.

We claim that this case can also be analyzed, and thus the universality problem for constraint sets including a bound on the vertex degree is decidable.

Proposition 3.4. *For constraint sets $\mathcal{C}$ including some star, the problem of the existence of a universal countable $\mathcal{C}$-free graph is decidable. A weakly universal $\mathcal{C}$-free graph will exist if and only if one of the following conditions holds:*

1. *$\mathcal{C}$ contains a path;*

2. *$\mathcal{C}$ contains a generalized 3-star $S(k_1, k_2, k_3)$ consisting of a central vertex v_0 and paths P_i of length k_i for $i = 1, 2, 3$ attached to v_0. In addi-*

tion, any maximal infinite connected $\mathcal{C}$-free graph is almost periodic, in a sense explained below.

At this point we will give just a sketch of the structural analysis, and put more about that in an appendix. The issue of decidability involves making some estimates explicit (which is not problematic), but also one must make explicit the analysis of a set of infinite words constructed from a particular finite set of finite words. We have convinced ourselves that this is manageable, but the reader is welcome to draw his own conclusions.

We now explain the particular notion of almost periodicity we have in mind here, which is very concrete. Note however that any reasonable notion of almost periodicity with respect to finite data in condition (2) would force the number of graphs under consideration to be countable, and thus imply that a universal one exists.

Definition 3.5.

1. Let G be a finite connected graph with two specified base points u_1, u_2. We let $G^{\mathbb{Z}}$ denote the graph obtained from the disjoint union of copies G_i of G ($i \in \mathbb{Z}$) by identifying the vertex u_2 of G_i with the vertex u_1 of G_{i+1}. The subgraph $G^{\mathbb{N}}$ is constructed in the same way from copies of G indexed by $\mathbb{N}$.

2. A graph H will be called *almost periodic* if it is periodic, or can be obtained from a periodic graph of type $G^{\mathbb{N}}$ by attaching one more finite graph G' with base point u to G_0, by identifying the base point u in G' with u_1 in G_0. Equivalently, H is either of the form $G^{\mathbb{Z}}$ or is obtained from a graph of the form $G^{\mathbb{N}}$ by adjoining finitely many vertices and edges, since any fixed finite initial segment G^n of $G^{\mathbb{N}}$ can be treated as part of the one additional graph G'.

As defined, our periodic graphs are connected, and the base points are cut points. In particular the blocks are finite of bounded order, and contained in the finite graph G taken as the initial building block. Much the same applies to almost periodic graphs. Clearly these can be construed as coded by words in a finite alphabet, but in a particularly simple way.

Proof of Proposition 3.4, sketch, cf. Appendix 1, §5. The case in which the set $\mathcal{C}$ contains a path was treated in Lemma 3.2, and the case in which $\mathcal{C}$ contains no path and no generalized star S was treated in Proposition 3.3. So we are left with the case in which $\mathcal{C}$ contains no path, but does contain some generalized star S.

Let G be $\mathcal{C}$-free. As G omits S, we argue first that the blocks of G have bounded diameter. As G has bounded vertex degree, it then follows that the blocks of G have bounded order.

Again, as G omits S, the underlying tree structure T on the blocks is path-like: there is a path P in T, with or without an end point, such that the remainder of T decomposes into connected components of bounded size. P is almost unique, apart from the first few vertices in the case where P has an endpoint.

We can define P more carefully by defining an appropriate set A of cut points v of G intrinsically, in terms of the sizes of connected components of $G \setminus \{v\}$; we require two large connected components. Then A inherits the graph structure of P; blocks of G which contain two vertices of A correspond to points of P between successive points of A. The connected components of $G \setminus A$ may be viewed as attached to one or two vertices of A; those that meet a block containing two successive vertices of A are viewed as attached to those two vertices, while for the other components there will be a unique vertex of A linked to the component by an edge.

Now orient the path P, taking the natural orientation if P has an endpoint, or an arbitrary orientation, otherwise. Then we can associate to the vertex $a \in P$, its successor b, and the set of components of $G \setminus A$ attached either to a or to a, b. Take the union of these components together with the vertices a, b, and let G_a be the induced graph on this set, with the vertices $u_1 = a$ and $u_2 = b$ taken as base points.

G can be considered as the connected sum of the G_a ($a \in A$) of order type $\mathbb{N}$ or $\mathbb{Z}$ and can be associated with the infinite word W whose successive terms are the isomorphism types of the structures (G_a, a, b). This is a finite alphabet, so there will be a long word w which repeats in W, giving a contiguous subword of W of the form $ww'w$. We consider the periodic word $(ww')^{\mathbb{Z}}$, and the corresponding periodic graph G^* which is constructed from the finite graph associated with the word ww'.

As the condition that G is $\mathcal{C}$-free is a strictly local condition, involving subgraphs of G of bounded diameter, G^* inherits this condition as long as w and w' are sufficiently long.

We may choose the word w to occur infinitely often in W. If G is not almost periodic of type ww', then there is another word w'' for which $ww''w$ occurs in the word W and $(ww'')^{\mathbb{Z}}$ is not a shift of $(ww')^{\mathbb{Z}}$. Then taking products of powers of (ww') and (ww'') we get $2^{\aleph_0}$ words corresponding to $2^{\aleph_0}$ nonisomorphic maximal connected $\mathcal{C}$-free graphs.

□

If we ask for strong universality we arrive at much more restrictive conditions: the infinite connected components must be trees, as otherwise a single maximal $\mathcal{C}$-free graph can be varied by taking a subgraph containing a spanning tree, and if necessary adjoining additional trees to obtain a $\mathcal{C}$-free graph which is maximal with respect to embeddings as an induced subgraph. But we prefer to turn now toward the general theory.

We will see shortly that much of the foregoing analysis works perfectly well in general, with *no bound on the vertex degree,* if (a) one confines oneself to the strongly universal case and (b) one replaces the straightforward notion of *connected component* (which is relevant only in the case of bounded degree), by a more delicate notion whose precise interpretation depends on the particular constraint set $\mathcal{C}$ under consideration, and which reduces to the connected component in the bounded degree case. It is only in this more general setting of unbounded degree that the model theoretic point of view becomes relevant. First we consider some additional examples illustrating the boundary between existence and nonexistence of universal graphs.

Fact 3.6.

1. *[17] Let C be a 2-connected graph. Then there is a countable universal C-free graph if and only if C is complete.*

2. *[12] Let $\mathcal{C}$ be a finite set of cycles. Then there is a countable universal $\mathcal{C}$-free graph if and only if $\mathcal{C}$ consists of all the odd cycles up to some fixed size.*

In general, the way to analyze the class $\mathcal{Q}_{\mathcal{C}}$ of $\mathcal{C}$-free graphs with respect to Problem B is the following. One associates to the class $\mathcal{C}$ in a very direct way a notion of $\mathcal{C}$*-algebraic closure*; for each set of vertices A in a $\mathcal{C}$-free graph G, this gives us a set $\mathrm{acl}_{\mathcal{C}}(A)$ containing A. At the outset one may take the following definition, which eventually will need to be made far more explicit: the vertex $v \in G$ is in $\mathrm{acl}_{\mathcal{C}}(A)$ if for any $\mathcal{C}$-free graph G^* containing A as an induced subgraph, the set of all possible images of v under embeddings of G into G^* over A is a finite set. Later we will make more explicit the kind of information needed in G to pin down v in this way. But for the moment this definition will suffice.

3.2. Algebraic closure

Consider a few examples. If $\mathcal{C} = \emptyset$ then evidently $\mathrm{acl}_{\mathcal{C}}(A) = A$ for any set A. Indeed, just take G^* to be the amalgam of infinitely many copies of G over A. On the other hand, in the case of bounded vertex degree with which we began, it is clear that $\mathrm{acl}_{\mathcal{C}}(a)$ is the connected component of a, and that for any set A we have

$$\mathrm{acl}_{\mathcal{C}}(A) = \bigcup_{a \in A} \mathrm{acl}_{\mathcal{C}}(a)$$

This last condition, which is weaker than degeneracy, will be called *unarity* here.

The property of most interest in this context will be *local finiteness.* The operator $\mathrm{acl}_{\mathcal{C}}$ will be said to be locally finite if $\mathrm{acl}_{\mathcal{C}}(A)$ is finite whenever A is. The following general result brings us to the heart of the matter and gives us a criterion for universality which can be made both explicit and purely combinatorial.

Theorem 2 ([11])**.** *Let $\mathcal{C}$ be a finite set of connected finite graphs. Suppose that* $\mathrm{acl}_{\mathcal{C}}(\cdot)$ *is a locally finite operator. Then there is a universal $\mathcal{C}$-free graph.*

The proof shows that in this case there is a canonical universal $\mathcal{C}$-free graph. It can be described as follows.

Definition 3.7. A $\mathcal{C}$-free graph G is *strongly universal* if for any finite subset A of G and any countable $\mathcal{C}$-free graph G^* containing G, there is an embedding of G^* into G over A.

Note that a strongly universal $\mathcal{C}$-free graph G is universal: if G_1 is $\mathcal{C}$-free then take $A = \emptyset$ and let G^* be the disjoint union of G and G_1. Furthermore there is at most one strongly universal countable $\mathcal{C}$-free graph, up to isomorphism, by a back-and-forth argument.

We could generalize this theorem to give an exact characterization of constraint sets $\mathcal{C}$ allowing a universal countable $\mathcal{C}$-free graph. But more progress comes from the theorem as we have stated it, because the local finiteness condition is much easier to work with, and because the exceptional cases where the local finiteness condition fails but a universal graph exists can be treated on an ad hoc basis, in a second round of analysis. We will see this more concretely when we discuss the analysis of tree constraints.

To tie up the knot on what we have said so far: in the case of graphs of bounded degree, the local finiteness condition says that connected $\mathcal{C}$-free graphs are finite. Thus for graphs of bounded degree the theorem is obvious, but it takes a certain body of theory to prove the theorem in general.

Theorem 2 needs to be supplemented by a close study of the operator $\mathrm{acl}_{\mathcal{C}}$ and how it is determined by $\mathcal{C}$. A good place to begin is the degenerate case.

Lemma 3.8 ([11, Lemma 5 and Theorem 4])**.** *Let $\mathcal{C}$ be a finite set of finite connected graphs. Then the following conditions are equivalent.*

1. $\mathrm{acl}_{\mathcal{C}}$ *is degenerate, that is:* $\mathrm{acl}_{\mathcal{C}}(A) = A$ *for all A;*

2. *$\mathcal{C}$ is closed under homomorphism in the following sense: for $C \in \mathcal{C}$ and for $\bar{C}$ a homomorphic image of C, there is $C' \in \mathcal{C}$ which embeds in $\bar{C}$.*

Homomorphisms between graphs are functions carrying vertices to vertices so as to induce a map from edges to edges. In particular homomorphisms do not identify two adjacent vertices, because the edge between them would go to a loop, and our formalism excludes loops.

One might rephrase the homomorphism condition as follows: a homomorphic image of a forbidden graph is forbidden.

Proof. ($2 \implies 1$):

We suppose (2): no $\mathcal{C}$-free graph is a homomorphic image of a graph which is not $\mathcal{C}$-free.

Suppose $A \subseteq G$, a $\mathcal{C}$-free graph, and $v \in G \setminus A$. Let G^* be the amalgam over A of infinitely many copies G_i of G. Then G is a homomorphic image of G^* and thus G^* is $\mathcal{C}$-free. But there are infinitely many images of v over A in G^* and thus $v \notin \mathrm{acl}_{\mathcal{C}}(A)$. So (1) holds.

($1 \implies 2$):

Any homomorphism can be obtained by composing two kinds of maps $f : G_1 \to G_2$: isomorphisms from G_1 to a subgraph (not necessarily an induced subgraph) of G_2, and maps in which G_2 is the result of identifying two vertices of G_1. Since the image of a forbidden subgraph under an embedding is forbidden, only maps of the second kind need concern us.

Suppose that we have $C \in \mathcal{C}$ and u, v vertices of C so that the graph $\bar{C}$ obtained from C by identifying u and v to a single vertex $\bar{u}$ is $\mathcal{C}$-free. Let $A = \bar{C} \setminus \{\bar{u}\}$. Then we claim that (1) is violated and specifically $\bar{u} \in \mathrm{acl}_{\mathcal{C}}(A)$. Indeed under any embedding of $\bar{C}$ into a $\mathcal{C}$-free graph G^*, the image of $\bar{u}$ is determined by the image of A, since two distinct images u_1, v_1 would allow us to reconstruct C in G_1. $\square$

This already goes some distance to explaining Fact 3.6. For a single constraint C, the operator acl_C is degenerate if and only if C is complete. For a finite set of cycles $\mathcal{C}$, the operator $\mathrm{acl}_{\mathcal{C}}$ is degenerate if and only if it consists of odd cycles up to some fixed size. Thus all of the positive cases covered by Fact 3.6 follow from the degenerate case.

On the negative side, the nondegeneracy of $\mathrm{acl}_{\mathcal{C}}$ is certainly not adequate to refute the existence of a universal $\mathcal{C}$-free graph, but in the two cases covered by Fact 3.6, what is needed is a two-stage process in which first, specific examples are constructed showing that $\mathrm{acl}_{\mathcal{C}}$ is not locally finite, and secondly, the construction is shown to have the capacity to incorporate enough latitude that it can be varied in $2^{\aleph_0}$ different ways, or in other words there are an infinite number of "free choices" which can be made during the construction. In the case of a 2-connected but incomplete constraint C this requires a very uniform construction which does not depend on the particular structure of C, while in the case of a set of cycles the situation is a good deal more concrete from the beginning, and it is just a matter of varying a construction given earlier for the case of a single cycle.

It does not seem that this final step, in which the construction is varied, can be usefully covered by a general theorem. On the other hand in all critical cases treated to date, the essential difficulty is overcome at the previous stage, when the dividing line between local finiteness and its failure is accurately identified. In practice, the absence of a theoretical path from the failure of local finiteness to the nonexistence of a universal graph has not been a major difficulty. And our thesis is that the essence of Problem $B_{\mathcal{Q}}$, at a combinatorial level, is captured by the following variation.

Problem (Problem $\tilde{B}_{\mathcal{Q}}$)**.** *Given the constraint set $\mathcal{C}$, determine whether* $\mathrm{acl}_{\mathcal{C}}$ *is locally finite.*

We will therefore focus on this problem. All the negative results to date have followed the route we have described, first refuting local finiteness, then exploiting the choices available to produce uncountably many incompatible structures, all of the form $\mathrm{acl}_{\mathcal{C}}(A)$ for some fixed finite A. We will discuss more such cases below, in which a single constraint is involved: namely, the case of 2-bouquet constraints, and the case of tree constraints. We also still need to give the promised analysis of $\mathrm{acl}_{\mathcal{C}}$ which is essential to the systematic investigation of Problem $\tilde{B}$, and we will come to that shortly.

But let us return for a moment to the special case of graphs of bounded vertex degree, and consider Problem $\tilde{B}$ in that context. This amounts to replacing "algebraic closure" by connected component.

Problem (Problem $\tilde{B}_0$)**.** *Given a finite set $\mathcal{C}$ of connected finite graphs, determine whether the connected graphs in $\mathcal{Q}_\mathcal{C}$ are of uniformly bounded size.*

As it happens, this last problem is not at all a good model for the general case, and in fact this problem has a straightforward solution. On the one hand, if an infinite star or an infinite path is $\mathcal{C}$-free, then the answer is negative. On the other hand, if the constraint set $\mathcal{C}$ contains some star and also some path, then the answer is positive.

This may appear disconcerting if we wish to use the locally finite case as the basis of our intuition about $\mathrm{acl}_\mathcal{C}(\cdot)$, but nonetheless, there remain strong parallels between the general notion of algebraic closure and the notion of connected component which are worth developing, along the following lines:

We naturally think of the connected component of a as "generated" by the adjacency relation, or explicitly:

$$\mathrm{acl}_\mathcal{C}(A) = \bigcup_n \Delta^n(A)$$

where $\Delta^n(A) = \{b : d(A, b) \leq n\}$, the n-th cumulative iterate of the operator

$$\Delta(A) = \{b : d(A, b) \leq 1\}$$

In the bounded degree case, where $\mathrm{acl}_\mathcal{C}(a)$ is the connected component of a, the operator Δ has the following desirable properties: $\Delta(A)$ is defined explicitly and concretely; Δ and each of its iterates is locally finite; and $\mathrm{acl}_\mathcal{C}(A)$ is exhausted by the iterates of Δ. In particular our problem $\tilde{B}_0$ is a problem about the length of this iteration (i.e., the diameter of the graph).

All of this goes over to the general case, when our graphs are not assumed to have bounded vertex degree, except that the definition of the associated operator Δ will now depend in general on the choice of $\mathcal{C}$, and will be denoted $\Delta_\mathcal{C}$. However, the solution of Problem $\tilde{B}_0$ does not translate into a solution of Problem $\tilde{B}$—or if it does, it requires more subtlety than anything we have tried.

3.3. Immediate algebraic closure

We wish now to describe an operator $\Delta_\mathcal{C}(A)$ associated with a finite set $\mathcal{C}$ of connected finite graphs, with the following properties.

1. $\Delta_{\mathcal{C}}(A)$ is finite if A is finite;

2. $\mathrm{acl}_{\mathcal{C}}(A) = \bigcup_n \Delta^n_{\mathcal{C}}(A)$;

3. $\Delta_{\mathcal{C}}(A)$ is directly and explicitly determined by $\mathcal{C}$.

We will say in this case that the elements of $\Delta_{\mathcal{C}}(A)$ are "immediately algebraic"' over A.

The required definition goes as follows. We follow [11], with some minor modifications. What is needed is mainly a notion of freeness of one finite subgraph over another, but we need also a subsidiary notion of richness for this notion to correlate properly with algebraic closure.

Definition 3.9. Fix a finite constraint set $\mathcal{C}$ as usual, and a $\mathcal{C}$-free graph G.

1. If $X \subseteq Y$ are graphs with X an induced subgraph of Y, denote by Y^∞_X the graph formed by amalgamating infinitely many copies of Y over the subgraph X; define Y^n_X similarly for n finite. If in addition Y is an induced subgraph of G, we say that Y is *free* over X in G if Y^∞_X embeds in G as an induced subgraph.

2. We say that G is *$\mathcal{C}$-rich* if for every pair of finite induced subgraphs $X \subseteq Y$ in G, if Y^∞_X embeds into some $\mathcal{C}$-free graph G^* containing G, then Y^∞_X embeds into G.

3. Let $A \subseteq X \subseteq Y$ be graphs, each an induced subgraph of G, with Y free over X in G. We say that X is a *base* for Y over A if Y is free over X, but is not free over any proper subset of X containing A.

4. Let A be an induced subgraph of G. We say that the vertex $v \in G$ is *immediately algebraic* over A if there are induced subgraphs $A_0 \subseteq X \subseteq Y \subseteq G$ with:

   (a) $A_0 \subseteq A$;

   (b) $v \in X$;

   (c) Y embeds as a subgraph in some $C \in \mathcal{C}$, and $|Y| < |C|$;

   (d) X is a base for Y over A_0.

5. $\Delta_{\mathcal{C}}(A)$ is the set of immediately algebraic elements of G over A.

It is not hard to show that every countable $\mathcal{C}$-free graph G embeds in a countable $\mathcal{C}$-free and $\mathcal{C}$-rich graph. Now we verify that this notion of immediate algebraicity meets our requirements, if the ambient graph G is $\mathcal{C}$-rich.

Proposition 3.10. *With $\mathcal{C}$ a finite constraint set consisting of connected finite graphs, and with G a fixed $\mathcal{C}$-free and $\mathcal{C}$-rich graph, the following hold.*

1. $\Delta_{\mathcal{C}}(A) \subseteq \mathrm{acl}_{\mathcal{C}}(A)$;
2. $\Delta_{\mathcal{C}}(A)$ *is finite if* A *is;*
3. $\mathrm{acl}_{\mathcal{C}}(A) = \bigcup_n \Delta^n_{\mathcal{C}}(A)$.

Proof.

1. If v is immediately algebraic over A then we fix witnesses A_0, X, Y in G as in the definition. Then we claim $v \in \mathrm{acl}_{\mathcal{C}}(A_0)$.

Assuming the contrary, we have G^* $\mathcal{C}$-free containing G, and infinitely many distinct images v_i of v under embeddings f_i of G into G^* over A. We consider the sets $X_i = f_i[X]$. Now apply the Δ-system lemma to the collection of finite sets X_i of fixed size. Then restricting to an infinite subset of the given embeddings, we may suppose that there is a fixed set $\bar{A}$ such that

$$X_i \cap X_j = \bar{A}$$

for all $i \neq j$. The inverse images $f_i^{-1}(\bar{A})$ are subsets of X; we may assume these sets coincide.

As Y is free over X in G, for each i we can find a copy Y'_i of Y in G, so that the images $Y_i = f_i[Y']$ also satisfy

$$Y_i \cap Y_j = \bar{A}$$

for $i \neq j$. Thus the Y_i are free over $\bar{A}$ in G^* and hence by $\mathcal{C}$-richness, Y is free over $f_i^{-1}(\bar{A})$ in G. Then by the minimality of X, we have $f_i^{-1}(\bar{A}) = X$ for all i, and in particular $v_i \in \bar{A}$. So these elements cannot be distinct.

2. We simply repeat the proof of (1).

Suppose that $\Delta_{\mathcal{C}}(A)$ is infinite, and for each $v \in \Delta_{\mathcal{C}}(A)$ pick a corresponding "witness" (A_v, X_v, Y_v) according to the definition. We may suppose $A_v = A$ for all v, and that the isomorphism type of the quadruple (A, X_v, Y_v) is fixed. By the Δ-system lemma we may suppose once more that we have a set $\bar{A}$ such that

$$X_u \cap X_v = \bar{A}$$

for all u, v distinct, and again by freeness of Y_v we may suppose that $Y_u \cap Y_v = \bar{A}$ as well. So again, $Y = Y_u$ is free over $\bar{A}$ and by minimality $\bar{A} = X_u$ for all u, and $u \in \bar{A}$, a contradiction.

3. It is easy to see that $\mathrm{acl}_{\mathcal{C}}(\mathrm{acl}_{\mathcal{C}}(A)) = \mathrm{acl}_{\mathcal{C}}(A)$ and thus the inclusion $\Delta^n_{\mathcal{C}}(A) \subseteq \mathrm{acl}_{\mathcal{C}}(A)$ follows from (1). The other direction is less formal.

Set $\hat{A} = \bigcup_n \Delta^n_{\mathcal{C}}(A)$. Consider the graph $G^*_{\hat{A}}$. It suffices to show that this graph is $\mathcal{C}$-free, as it allows infinitely many embeddings of G disjoint over $\hat{A}$, forcing $\mathrm{acl}_{\mathcal{C}}(A) \subseteq \hat{A}$.

So suppose toward a contradiction that some $C \in \mathcal{C}$ embeds into $G^*_{\hat{A}}$. Consider $\hat{A}_0 = C \cap \hat{A}$. The graph $G^*_{\hat{A}}$ is the union of copies G_i of G, with A in common; let Y_i' be $C \cap G_i$ and let Y_i be the corresponding induced subgraph of G itself. Here we need only concern ourselves with the finitely many graphs Y_i for which $Y_i \neq \hat{A}_0$.

If each of these graphs Y_i is free over $\hat{A}_0$ in G, then we may choose copies Z_i of Y_i in G which are pairwise disjoint over $\hat{A}_0$, getting an embedding of C into G, and a contradiction. So some Y_i is not free over $\hat{A}_0$ in G. But Y_i is free over Y_i, vacuously, and hence there is a base X for Y_i over $\hat{A}_0$, and this base properly contains $\hat{A}_0$. But as $X \subseteq \Delta_{\mathcal{C}}(\hat{A}_0) \subseteq \hat{A}$, we find $X \subseteq Y_i \cap \hat{A} = \hat{A}_0$, a contradiction. $\square$

The concluding portion of the last proof gives a more detailed indication of which subgraphs $Y \subseteq C$ are actually relevant, namely those which are part of a collection of subgraphs of C which could be amalgamated over a common part to give C. This is a useful bit of information in practice. The same analysis can be used to clarify the property of unarity, as follows.

Fact 3.11 ([11, Proposition 6 and remarks following]). *For C a single connected finite constraint, the following are equivalent.*

1. $\mathrm{acl}_{\mathcal{C}}$ *is unary;*

2. *The blocks of C are complete graphs.*

The implication (2) $\implies$ (1) was given more generally for finite constraint sets in [11], and the converse direction, which holds for single constraints, was only mentioned, without proof. One can read [17] as exploiting this principle, though this formalism is not used there.

When the operator $\mathrm{acl}_{\mathcal{C}}$ is unary, the simplifications resulting from the restriction to a consideration of $\Delta_{\mathcal{C}}(a)$ and its iterates can be substantial. Conversely, the following conjecture was stated in [10].

Conjecture 3. *When $\mathcal{C}$ contains a single constraint, and the operator* $\mathrm{acl}_{\mathcal{C}}$ *is not unary, no universal $\mathcal{C}$-free graph will exist.*

In [10] we also mentioned the possibility that the same would hold for any finite constraint set, but prudently refrained from giving that as a conjecture. We will refute that more general form in Appendix 2 (§6). Given that refutation, there is no theoretical basis for the foregoing conjecture, but it corresponds to our sense that the case of a single constraint should be manageable by explicit analysis. The methods used in [17] and [12], which exploit failures of unarity, are suggestive. Furthermore an idea of Shelah which we call *pruning* allows an inductive approach to such problems. We take this up next.

3.4. Pruning

One very simple idea of Shelah has had a substantial impact and is far from exhausted: the effect of "pruning" on universality problems. This was introduced and applied in [10].

One obstruction to a clean theory has been the circumstance that a "tighter" set of constraints does not necessarily yield a "simpler" class as far as the problem of the existence of universal graphs is concerned. Indeed, if we have no constraints at all then there is a universal graph: the Rado graph, which here falls under Theorem 2 via Lemma 3.8, a clear case of overkill. What seems more to the point is an example which turns up in an analysis of 2-bouquets given in [14]. We write $K_m \wedge K_n$ for the graph with two complete blocks which are complete graphs of order m and n; in other words, K_m and K_n joined at a common point. One then has the following:

Fact 3.12 ([14])**.** *Let $C = K_m \wedge K_n$ be a 2-bouquet. Then there is a universal C-free graph if and only if the following conditions are satisfied:*

1. $\min(m, n) \leq 5$;
2. $(m, n) \neq (5, 5)$.

In particular we have universal C-free graphs for $K_4 \wedge K_5$ and $K_6 \wedge K_5$ but not for $K_5 \wedge K_5$. One can make some sense of this by viewing the "symmetric" and "asymmetric" cases as slightly out of phase with each other, but the main point is that this type of result strongly suggests that the analysis of one case may not cast much light on any other.

Fortunately, that suggestion is wrong, and the following comes as a welcome surprise.

Lemma 3.13 ([10, Proposition 2.3]). *Let T be a tree for which there is a countable universal T-free graph, and let T' be the tree obtained from T by removing all its leaves (pruning). Then there is a universal T'-free graph.*

This was exploited, with some labor, to confirm a long-standing conjecture as to which trees T do correspond to universal graphs [10]. As the list of such trees is very short, Lemma 3.13 reduces the number of essentially distinct cases which need to be analyzed to a manageable size.

But pruning can be applied in a very general form, and provides a powerful point of departure for future analyses. We give the general statement.

If $\mathcal{C}$ is a finite set of connected finite graphs, we may view each $C \in \mathcal{C}$ as made up of a tree of blocks. The blocks which occur as leaves in such a tree decomposition will be called *block-leaves* of $\mathcal{C}$; these are really *pointed blocks* (v, B) with v a vertex in B representing its point of attachment to the rest of C. A *minimal block-leaf* (v, B) is one for which an embedding $(v', B') \to (v, B)$ as a subgraph is necessarily an isomorphism, for any other block-leaf (v', B') of a graph in $\mathcal{C}$.

If (v, B) is a block-leaf of the graph C, pruning C at (v, B) means removing $B \setminus \{v\}$. Pruning C (globally) with respect to a block-leaf (v, B) means pruning C at (v', B') for each block-leaf of C which embeds as a subgraph into (v, B). Pruning a set $\mathcal{C}$ of finite graphs with respect to the pointed graph (v, B) (where B is 2-connected) means pruning each graph $C \in \mathcal{C}$ with respect to (v, B). For example, if $\mathcal{C}$ is a set of trees then its blocks are of order 2 and one prunes the set $\mathcal{C}$ by removing the leaves from each tree.

Lemma 3.14 ([10, Proposition 2.3]). *Let $\mathcal{C}$ be a finite set of connected finite graphs, and (v, B) a pointed 2-connected graph. Let $\mathcal{C}' = \{C' : C \in \mathcal{C}\}$ be the result of pruning $\mathcal{C}$ with respect to (v, B) (pruning at all occurrences of subgraphs of (v, B)). If there is a universal countable $\mathcal{C}$-free graph, then there is a universal countable $\mathcal{C}'$-free graph.*

So there is a natural inductive approach which is likely to be part of any very direct attack on the universality problem.

Proof. We first define an anti-pruning operation: for any graph G, let G^* be the result of freely attaching infinitely many copies of (v, B) to each vertex of G, taking v as the point of attachment.

If G is $\mathcal{C}'$-free then we claim that G^* is $\mathcal{C}$-free. Suppose that $C \in \mathcal{C}$ embeds into G^*. Any block of C whose image contains a vertex outside G will lie in one of the newly adjoined copies of (v, B). And in view of the structure of G^*, it will necessarily be a block-leaf of C. So C' itself is forced into G, a contradiction.

Now suppose that G is countable universal $\mathcal{C}$-free and let G_0 be the induced subgraph on the set of vertices $u \in G$ such that there are infinitely many copies of (v, B) with common vertex $v = u$, and otherwise disjoint. We claim that G_0 is universal $\mathcal{C}'$-free.

Certainly G_0 is $\mathcal{C}'$-free: if $C \in \mathcal{C}$ and C' embeds into G_0, then the definition of G_0 gives an extension of this embedding to an embedding of C into G, and a contradiction.

So now suppose H is countable $\mathcal{C}'$-free. Embed H^* into G; then the vertices of H are carried into vertices of G_0, and our claim follows. $\square$

We feel that this line of attack is very promising and may eventually lead to a complete solution to Problem B for the case of a single constraint. To put the matter formally:

Conjecture 4. *The existence of a universal C-free graph is a decidable problem, for C a single finite connected forbidden graph.*

In this particular case, what we have in mind is something close enough to an explicit solution to trivialize the decidability question.

Indeed, we felt for some time that the general case might be, roughly speaking, a combination of the 2-connected case cited in Fact 3.6 and the case of trees, which runs as follows.

Fact 3.15 ([10, 14])**.** *Let T be a finite tree. Then the following are equivalent.*

1. *There is a universal T-free graph;*
2. *T is either a path or can be obtained from a path by adding one additional vertex and one corresponding edge.*

We are hopeful that the 2-connected case goes over quite generally: that a constraint C for which there is a countable universal C-free graph must have complete blocks (Conjecture 3, Fact 3.11). On the other hand we will not say that we expect the underlying tree structure to be *quite* as simple as it is when the constraints are actually trees—something suggested by the optimistic [10, Conjecture 2]—but we think it is plausible that the "path or near-path" rule will be the main case, with some more limited examples of other types, notably the following.

Conjecture 5 (Hairy Ball Conjecture)**.** *Let C be a graph obtained from a complete graph K by attaching at most one path to each vertex. Then the operator acl_C is locally finite, and in particular there is a universal C-free graph.*

This last conjecture can be decoded to a completely explicit graph theoretic problem, and we will work that out below.

In short, the general form of the answer to the universality problem for the case of a single constraint is still not quite in sight; but the tools for pinning it down in that case appear to be in hand.

We conclude this subsection with some additional comments on the pruning construction and its application via Lemma 3.14.

We have some freedom in general to choose the pointed graph which determines the pruning chosen. This graph should of course contain a block-leaf actually occurring in one of the members of $\mathcal{C}$, as otherwise $\mathcal{C}' = \mathcal{C}$. When pruning sets of trees there is only one possible type of block, so there is only one type of pruning in that case.

One can prune even more generally: the essential property of a block-leaf (v, B) of a graph C is that $B \setminus \{v\}$ is a connected component of $C \setminus \{v\}$. To date the most useful kind of pruning has been pruning with respect to a minimal block-leaf but Shelah has given good reasons in unpublished notes to expect that the more general pruning operation will be useful in practice, at a later stage of analysis.

3.5. The Hairy Ball Constraint

In this section we consider a constraint C consisting of a complete graph K of order n together with a single finite path P_v with endpoint v (possibly of length 0) attached to each vertex v of K .

Of particular importance will be the subgraphs C_v of C associated with vertices v of C as follows: let C'_v be the connected component of $C \setminus \{v\}$ containing $K \setminus \{v\}$, and let C_v be the induced graph on $C'_v \cup \{v\}$. Also, let P_v be the path from v supplementary to C_v, so that $C = C_v \cup P_v$.

In order to state our conjecture in concrete terms we make use of the analysis of "acl" undertaken in Definition 3.9, using the notion "free over" a subgraph.

Conjecture 6 (Hairy Ball II)**.** *Let G be a graph containing a sequence of vertices $(v_i)_{i \in \mathbb{Z}}$ such that the v_i lie along a 2-way infinite path Q which is free* *in G over the $(v_i)_{i\in\mathbb{Z}}$, and each vertex v_i belongs to a subgraph C_i so that the pair (v_i, C_i) is isomorphic to a pair (v, C_v) in the notation above. Then C is isomorphic to a subgraph of G.*

In this formulation, we intend the vertices v_i to be enumerated in their order along Q, with respect to some orientation of Q. In general the vertices v_i will not exhaust the vertices of Q, but then the freeness condition allows us to take Q so as to avoid clashes between the remaining vertices of Q, and other vertices which may come into consideration as part of the structure of the graph in neighborhoods of the vertices v_i. We would expect that a proof in the case in which the v_i do exhaust the path Q would lead quickly to a proof in general.

Proposition 3.16. *Conjectures 5 and 6 are equivalent.*

Suppose first that Conjecture 6 fails, so that there is a configuration (v_i, C_i) inside a C-free graph G with the stated properties. It suffices to show that for each i there is $j > i$ so that $v_j \in \mathrm{acl}(v_i)$.

Now $(v_i, C_i) \cong (v, C_v)$ for some $v \in C$. Let Q be a path containing the sequence $(v_i)_{i\in\mathbb{Z}}$ and free over it. Let Q_i be a segment of Q of the same length as the supplement P_v to C_v in C, beginning at v_i, and in the positive direction along Q. As G is C-free, Q_i is not free over v_i, so let B be a base for Q_i over v_i. As Q is free over $(v_i)_{i\in\mathbb{Z}}$, B is a subset of $\{v_j : j > i\}$. Therefore any element of B is a vertex v_j with $j > i$ which lies in $\mathrm{acl}(v_i)$, as claimed.

To argue in the converse direction we will have to analyze the algebraic closure operation again in the manner of §3.3, but more explicitly; this was done in the general case for graphs with complete blocks in [11].

We use the following concrete notion of immediate algebraic closure.

Definition 3.17. $v' \in \mathrm{acl}'(v)$ if the following holds:

There is a path P_v with endpoint v, not free over v, and a base B for P_v over v such that v' is the nearest vertex to v in B.

Lemma 3.18. *For G C-rich and C-free,* acl' *generates* acl_C*.*

Proof. We saw in Proposition 3.10 that acl' is contained in acl. So taking A not algebraically closed, and finite, we must find $v \in A$ with $\mathrm{acl}'(v)$ not contained in A.

The infinite amalgam G_A^∞ is not C-free, so embed C into G_A^∞; we will now identify C with its image in G_A^∞. Call $u \in C$ a *transition point* if there are distinct factors G_i, G_j of G_A^∞ such that u has neighbors in both $G_i \setminus A$ and $G_j \setminus A$. We may suppose that the embedding of C into G_A^∞ has been chosen to minimize the number of transition points. But since G is C-free, there must be at least one such point.

Choose a transition point $a \in C$ such that the path P_a outward from a in C has minimal length. Notice that $a \in C \cap A$ in view of the structure of G_A^∞. Then the neighbors of a in C lie in exactly two factors G_i, G_j; let G_i be the factor that contains the neighbors of A in C_a, and G_j the factor containing the neighbor of a in P_a.

If P_a is free over a in G, then we can replace P_a by a path in G_j and reduce the number of transition points. So P_a is not free over a in G. Let B be a base for P_a over a, and let a' be the closest vertex to a in B. If $a' \notin A$ then we have what we have been aiming for: $a' \in \mathrm{acl}'(a) \setminus A$.

If $a' \in A$ we have to look a little more. As the interval (a, a') in P_a is free over a, a', it can be replaced by an interval in G_i disjoint from $C_a \cup P_{a'}$. The effect of this is to replace the transition point a by a new transition point a', and the path P_a by the shorter path $P_{a'}$. Repeating this argument (or phrasing the initial minimization a little more precisely) we arrive eventually at our claim. □

Now we complete the proof of Proposition 3.16. Suppose that Conjecture 5 fails, and thus the operator acl_C is not locally finite; as this operator is unary (Fact 3.15), for some a the set $\mathrm{acl}_C(a)$ is infinite. Then if we define $\mathrm{acl}_C^{(n)}$ as the n-th iterate of acl', Lemma 3.18 and Proposition 3.10 (2) show that $\mathrm{acl}_C^{(n)} \neq \mathrm{acl}_C^{(n+1)}$ for all n. Fix N, choose $a_N \in \mathrm{acl}_C^{(N)}(a) \setminus \mathrm{acl}_C^{(N-1)}$, and then choose by downward induction elements a_i for $i < N$ with $a_i \in \mathrm{acl}^{(i)}(a)$ and $a_{i+1} \in \mathrm{acl}'(a_i)$. Then inductively, $a_i \notin \mathrm{acl}^{(i-1)}(a)$.

Now as $a_{i+1} \in \mathrm{acl}(a_i)$ we have, for each i, a path P_i with endpoint a_i, and a base B_i for P_i over a_i such that a_{i+1} is the nearest vertex to a_i in B_i. In particular, the path $Q_i = [a_i, a_{i+1}] \subseteq P_i$ is free over a_i, a_{i+1}. Furthermore the vertices $(a_i)_{0 \le i \le N}$ are distinct, since for $i < j$ we have $a_i \in \mathrm{acl}^{(i)}(a)$, $a_j \notin \mathrm{acl}^{(i)}(a)$. So the paths Q_i can be glued together to give a path Q from a_0 to a_N which is free over $(a_0, a_1, \ldots, a_N)$. And the definition of acl' gives us a graph $C_i = C_{a_i}$ associated to each vertex a_i, as required.

Since our initial choice of N is arbitrary, an application of the compactness theorem of logic, or of König's tree lemma, shows that we can extend this to a similar pattern of order type $\mathbb{Z}$, all in a C-free context: so the failure of Conjecture 5 entails the failure of Conjecture 6.

3.6. Forbidden substructures and forbidden induced substructures

So far, our discussion has focused single-mindedly on forbidden subgraphs. But one can consider similar problems for forbidden substructures in any combinatorial setting. The general theory applies as long as the structures fit into our framework as relational systems.

The classification of the homogeneous universal structures has been carried out for graphs, digraphs, colored partial orders, permutation patterns, and in other cases [32, 8, 51, 7, *inter alia*]. This could be taken as a point of departure for a more general study of universal structures. In the case of permutation patterns, it would be very desirable to have a general theory of universal structures; in some cases these provide a canonical infinite limit for the class of finite structures (generalizing the Fraïssé theory). Such limits have been considered in the permutation pattern literature, but not systematically.

While it would be very nice to import the sort of model theoretic machinery we have for graphs, that is not actually feasible. That particular version of the theory depends on our ability to form disjoint unions, and more generally disjoint unions over a common substructure. So to bring this theory over to this interesting case would require further foundational work (Problem 12, §4.2). The general model theoretic point of view is still relevant and should have some sensible interpretation in this setting.

But we may consider universality problems for classes of general structures determined by finitely many forbidden substructures (or in model theoretic terms: weak substructures). Then the theory developed for graphs applies very well. At the same time, it has been shown in [13] that in a straightforward sense, the problem of the existence of universal $\mathcal{C}$-free *structures* is no more complicated than the special case in which the structures are simply graphs equipped with a coloring of the vertices by two colors; and for all we know, there may well be a reduction of the general problem to universality problems for ordinary graphs without additional structure, but that point remains open.

It would also seem natural to consider universality problems with finitely many forbidden *induced substructures*. This is a broader problem: to forbid one *substructure* A, it suffices to forbid the finite set of *induced substructures* which contain A and have the same elements.

But this turns out to be a problem of a radically different character. In the first place, the theory sketched here in terms of $\mathrm{acl}_{\mathcal{C}}$ breaks down completely when the constraint set $\mathcal{C}$ is an arbitrary finite set of forbidden induced substructures. We will see in particular that any universality problem involving finitely many forbidden induced substructures can be transformed to an equivalent problem with $\mathrm{acl}_{\mathcal{C}}$ degenerate. Thus Theorem 2 becomes irreparably false at this level of generality, eliminating our most useful tool.

Furthermore, we will show that Problem B becomes undecidable when one takes as constraints a finite set of forbidden substructures.

Theorem (4, §7)**.** *The existence of a universal $\mathcal{C}$-free graph, with $\mathcal{C}$ an arbitrary finite set of finite connected induced subgraphs, is an undecidable problem.*

In the present section, we will give a weaker form of Theorem 4 which involves fewer coding issues, and makes a similar point. We give the proof of Theorem 4 in Appendix 3 (§7).

The breakdown of our general theory in the context of universality problems for forbidden induced subgraphs is illustrated by the following.

Example 2. *Let $\mathcal{Q}$ be any class of graphs, and let $\mathcal{Q}^*$ be the class of structures of the form $(V, E, \sim)$ satisfying the following conditions:*

1. *$G = (V, E)$ is a graph;*
2. *$\sim$ is a congruence on the graph G, that is an equivalence relation on V satisfying the law*

$$a \sim a', b \sim b', E(a, b) \implies E(a', b')$$

   with E the edge relation.

3. *$G/\sim$ is a graph in $\mathcal{Q}$ with the induced edge relation $E(\bar{a}, \bar{b}) \iff E(a, b)$.*

Then there is a countable universal G^* in $\mathcal{Q}^*$ if and only if there is a countable universal G in $\mathcal{Q}$, and if $\mathcal{Q}$ is determined by a finite set of forbidden subgraphs, then $\mathcal{Q}^*$ is determined by a finite set of forbidden induced substructures.

Notice that in the foregoing construction, we cannot constrain $\sim$ to be an equivalence relation using forbidden substructures.

On the other hand, the corresponding operation of acl in $\mathcal{Q}^*$ is degenerate, because each vertex may have an infinite equivalence class consisting of indistinguishable elements. So the operation acl no longer conveys anything at all in this setting (or else, it becomes an operation on equivalence classes rather than on elements).

This shows that our theory of algebraic closure has no bearing on such cases.

Next we would like to show that problems of this type become undecidable when forbidden induced substructures are considered. For this we use Hao Wang's unconstrained domino problem, shown undecidable by Berger [2].

Wang's problem is a tiling problem. Wang tiles (which he called *dominoes*) may be thought of as unit squares with colors along the edges, which are to be used to tile the plane $\mathbb{Z}^2$, with colors matching along adjacent edges.

We will find it convenient to set this up a little more generally. The tiles we use can be thought of as unit squares, with each tile carrying a single color, and with arbitrary horizontal and vertical matching rules, saying which pairs of tile colors may occur successively, in either the horizontal or the vertical direction. To convert a Wang tile set with color set C into one of our form, we construe a tile with edge colors c_E, c_N, c_W, c_S as a tile carrying the "color" (c_E, c_N, c_W, c_S), and we give as the matching rule the requirement that corresponding entries agree, i.e. the c_E entry in one tile equals the c_W in the next one horizontally, with a north-south match in the vertical direction.

Thus the undecidability of the Wang tiling problem yields the undecidability of our ostensibly more general problem, which is all we will need. However one may encode our more general tiles as Wang tiles as follows: from any tiling of the plane by unit squares with centers on the lattice $\mathbb{Z}^2$, we derive a tiling of the plane by unit squares centered on the shifted lattice $(\mathbb{Z}+\frac{1}{2})^2$, where each of the new squares overlaps with four of the original ones. The squares of the shifted tiling may be treated as dominoes if we assign to each edge of a new tile the ordered pair of colors of the two tiles in which that edge lies, and these dominoes satisfy the color matching condition.

The (unconstrained) tiling problem is to tile the plane with a given finite set of tiles and specified tiling rules. The corresponding decision problem was originally posed by Wang for the case of his dominoes, and shown undecidable by Berger by an encoding of Turing machine computations. In particular, the decision problem, whether it is possible to tile the plane with a specified set of tiles and tile constraints in our sense, is undecidable.

It is useful to take note of the following reduction of the tiling problem.

For any specified tile set (and tiling rules), one of the following occurs.

1. It is possible to tile the plane with (some of) the specified tile types, and observing the rules; or
2. For some finite n, there is no acceptable tiling of an $n \times n$ square.

Since the tiling problem is undecidable, we see that there is no way to compute a relevant "test" value of n from the tile set.

We wish to convert each tile set into a related class of structures, determined by finitely many forbidden induced substructures, in such a way that tile sets which can be used to tile the plane correspond to classes of structures for which there is no countable universal object, thus reducing the undecidability of the latter problem to a known result.

This involves the consideration of what one might call "nonstandard" tilings. The definition of a tiling in $\mathbb{Z}^2$ depends on the structure of $\mathbb{Z}$ with the successor relation, which defines the relations "right neighbor," "left neighbor," "next above," and "next below" in $\mathbb{Z}$. Given any set of tiles and tiling rules, and any two directed graphs A and B, we can define analogously what is meant by an admissible tiling of $A \times B$: we place a tile at each point of the Cartesian product, and whenever we have a pair of points (a, b) and (a', b) in Z^2 with an edge relation $E(a, a')$, or (a, b) and (a, b') with an edge relation $E(b, b')$, we impose the corresponding tiling rule. We do not really need anything as exotic as the general case: in practice we will want A and B to be oriented paths (finite or infinite) or oriented cycles (or at worst, disjoint unions of such graphs). For example, when A and B are finite paths, a tiling of $A \times B$ is just a partial tiling of the plane; when A and B are both cycles, a tiling of $A \times B$ encodes a periodic tiling of the plane (more explicitly, a doubly periodic tiling).

Our encoding of tiling problems proceeds as follows.

Let the tile set be specified, and suppose there are k tiles, arranged in some particular order as $(t_1, \ldots, t_k)$. We consider structures G equipped with the following:

A. An asymmetric relation E; so (G, E) is a directed graph without loops or multiple edges.

B. A series of binary relations $T_i(x, y)$ for $1 \leq i \leq k$.

C. A unary relation A on G.

We impose the following constraints on G.

1. The in-degrees and out-degrees of (G, E) are bounded by 1; in other words, the components are oriented paths or cycles;

2. For any pair x, y in G, one and only one of the relations $T_i(x, y)$ holds;

3. Taking the relations T_i to encode a tiling of G^2 and E to give the local structure on G^2, the tiling constraints are respected

Let us clarify the picture. By our first condition, the components of G are oriented cycles or paths. The induced structure on G^2 is given by two relations, the horizontal successor relation $S_H((a, b), (a', b')) \iff E(a, a')\&b = b'$ and the vertical successor relation $S_V((a, b), (a', b')) \iff a = a'\&E(b, b')$. In particular the union of these relations gives G^2 the structure of an oriented graph which is locally a grid, and the connected components of G^2 are the products $A \times B$ with A, B connected components of G.

To tile G^2 means to tile $A \times B$ for each pair of components A, B of G, in such a way that the given tiling rules are respected. In other words, we tile rectangles. The tiling rules specify the tile types that can be assigned to points of $A \times B$ which are connected by the horizontal or vertical successor relations, coming from the successor relations on the oriented paths A and B respectively.

Most of our constraints here are just "forbidden substructures," including half of the clause "one and only one" in the second constraint. However the requirement that at least one of the relations T_i should hold is a constraint on induced substructures, and not at all the sort of constraint that can be imposed using forbidden substructures.

The point of the construction is the following.

Proposition 3.19. *Let T be a set of tiles with a specified set of tiling rules, and let $\mathcal{Q}_T$ be the corresponding class of structures. Then there is a universal countable structure in $\mathcal{Q}_T$ if and only if there is no tiling of the plane by the tile set T respecting the rules.*

Proof. Suppose first that there is such a tiling. Then consider structures consisting of a single 2-way infinite oriented path P together with an appropriate tiling of P^2, with an arbitrary interpretation of the predicate A on P. There are uncountably many such structures, and at most countably many of them embed into a countable structure in $\mathcal{Q}_T$, since P will go over to a connected component of the image.

Now suppose there is no such tiling. Then by König's lemma, there is some finite bound n on the length of a path P for which P^2 can be tiled respecting the tiling rules, or to put the matter more directly: there is a bound on the sizes of connected components of (G, E) for $G \in \mathcal{Q}_T$. We use this fact to build a countable universal structure for the class.

Observe that as there is no tiling compatible with the rules, no connected component of a structure in $\mathcal{Q}_T$ contains any cycles; otherwise, if C were such a cycle, the tiling on $C \times C$ would give rise to a periodic tiling of the plane. So the components of such structures are oriented paths of bounded length; let L be the maximum length of a path P for which P^2 carries a compatible tiling.

We will call a countable structure G in $\mathcal{Q}_T$ *homogeneous universal* if it satisfies the following condition: For any finite substructure H_0 of G which is a union of connected components of G with respect to E, and any embedding $\iota : H_0 \to H_1$ with $H_1 \in \mathcal{Q}_T$ finite, and with H_0 a union of connected components in H_1, there is an embedding of H_1 over H_0 into G, with the image of H_1 a union of connected components in G.

We claim that homogeneous universal structures exist, and are, as the terminology would indicate, in fact universal. We claim further that there is only one such up to isomorphism, but we do not require that fact.

We remark that we are following well-trodden model theoretic lines here, associated with Fraïssé ([23]).

Existence. One builds G as a union of finite structures G_i such that each G_i is a union of connected components in the next. It then suffices to show that having built G_i for some i, and taking H_0 a union of some connected components of G_i, and $\iota : H_0 \to H_1$ a corresponding extension as in the definition, the structure G_{i+1} can be manufactured so that G_i remains a union of connected components of G_{i+1}, and H_1 embeds into G_{i+1} over H_0, also as a union of connected components.

Indeed, first take for G_{i+1} the disjoint union of G_i and H_1 over H_0, so that its connected components are those of H_0, those of G_0 disjoint from H_0, and those of H_1 disjoint from H_0. Extend the unary predicate A arbitrarily. The tilings are defined within G_i and H_1, so it remains to define appropriate tilings of $P \times Q$ and $Q \times P$ when P and Q are components of $G_i \setminus H_0$ and $H_1 \setminus H_0$ respectively. Now P and Q are paths of length at most L, and therefore there are tilings of $P \times Q$ and of $Q \times P$ which respect the tiling rules. So extend the relations T_i correspondingly.

Universality. For any countable H in $\mathcal{Q}_T$, write H as a union of finite substructures H_i with each H_i a union of connected components of H. Embed $H_i \to G$ inductively so that the image at each stage is a union of connected components of G, using the homogeneous universality of G to carry out the inductive step.

For the uniqueness of G argue similarly, interchanging the roles of H and G at each successive step. □

Corollary 3.20. *There is no effective procedure to determine whether a class $\mathcal{Q}_T$ as above contains a countable universal model.*

We give this corollary to make a point about the difference between forbidden substructures and forbidden induced substructures. Of course, we cannot yet rule out a more subtle encoding of tiling problems by universality problems with forbidden substructures, but the requirement that every vertex of a grid carry some tile is naturally expressed by a forbidden *induced* substructure.

Our construction has two defects which are addressed in Appendix 3, §7. First, we have used a variable language, which is not really consistent with the framework we set out initially. This can be addressed by limiting oneself to two types of tiles and then allowing more elaborate constraints as tiling rules (essentially, various patterns of two tiles are taken to represent a set of tiles).

But one would also like to show that one can do the whole construction in the language of graph theory, so that our universality problem for graphs becomes undecidable as soon as one allows forbidden induced subgraphs. This is Theorem 4 of §7.

The reduction to a finite language is straightforward (§§7.1–7.2). The reduction to the language of graphs is more delicate.

4. Open Problems, Notes

We will present some of the open problems touched on in §§2,3, and conclude with a few technical remarks relating to points not developed in the text.

4.1. WQO problems

So far in concrete cases of the WQO problem for classes of structures determined by finitely many constraints, the theory gives us a good idea of where we should look for antichains and what sort of antichains we should look for. It says less about what we should do when there are no antichains and we need to prove the WQO property. Generally we reach at some point for Kruskal's tree theorem. This suggests the following ill-defined problem.

Problem 1 (WQO techniques)**.** *To what extent can the structure theorems needed to prove WQO for specific classes of tournaments defined by forbidden subtournaments be subsumed under the type of analysis occurring in the proof of the graph minor theorem, or an analog of that analysis for the case of tournaments?*

In [9] there was a brief discussion of the tournaments known to be in Λ_2. We have a given a somewhat more systematic account of the origin of the known antichains here and it is worth revisiting Λ_2 with this in mind.

Problem 2. *Which antichains arising by known constructions (either the doubling construction (Notation 2.15) or an encoding from $\mathcal{Q}^{(c)}$ (§2.5)) lie in Λ_2? Do such antichains exhaust Λ_2 completely?*

Note that we do not have very good control of these constructions for $c > 2$ but we do have good control for $c = 2$, and for the case of two constraints (i.e., for Λ_2) that should really be enough.

Problem 3 (Paths with wqo vertex color sets)**.** *Can one identify the set Λ when $\mathcal{Q}$ is the collection of structures consisting of a finite oriented path P together with a coloring of the vertices of P using a fixed set of colors, and the set of colors carries a wqo?*

Here an embedding must preserve edges and the coloring c of P must be compatible with the embedding f in the sense that $c \leq f(c)$.

We looked at the case of paths with vertex colorings using a fixed finite set of colorings in §1, but in this area, finite colorings can often be replaced by colorings using wqo sets of colors, with the corresponding notion of embedding, as we have had occasion to see already. However we have not taken up this generalization, so we leave it as a problem.

In addition, the problem of the precise relationship between the class $\mathcal{Q}^{(c)}$ of paths with c colors and other classes like the class of finite tournaments has not been worked out in any systematic way. So one could explore this more systematically. The first step would be the following problem, stated as a conjecture in §2.5.

Problem 4. *Let I be an almost periodic antichain in $\mathcal{Q}^{(2)}$, and k fixed. For $P \in I$ let $T_k(I)$ be the corresponding tournament using the sequence $(N_{k,n})$ of §2.5 as the base. Show that after removal of finitely many terms, $(T_k(P) : P \in I)$ is an isolated, minimal antichain in the quasiorder of finite tournaments, whose associated ideal is effectively determined (and uniformly effective, relative to the data determining I).*

More generally:

Problem 5. *Determine the "natural" embeddings of $\mathcal{Q}^{(c)}$ into the class $\mathcal{Q}$ of finite tournaments.*

Problem 4 suggests the following, mentioned in [9] without the hypothesis of isolation.

Problem 6. *Is it true that for any isolated minimal antichain I of finite tournaments, there is a bound k so that each $T \in I$ is covered by k linear orders?*

In any case, one should try to determine the structure of the isolated minimal antichains satisfying this condition. Whether isolation is relevant here is unclear, but as we also conjecture that the isolated antichains are dense, we believe that this hypothesis is harmless; these should be the only ones whose study is relevant to Problem A.

For $c = 2$ the problem is to determine appropriate tournaments T_n corresponding to paths P_n, for an infinite set of indices n. We proposed one family of examples $N_{k,n}$ which arises by first encoding a k-colored linear order into a tournament, and then reversing the edges corresponding to the successor relation.

For $c > 2$ one gets a useful embedding, in fact several for each value of c, but not all minimal antichains go over into minimal antichains; as we saw, they must be minimal in a strong sense, with respect to homomorphisms between color sets.

Problem 7. *To what extent are isolated minimal antichains in $\mathcal{Q}$, for the case of tournaments, derived from $\mathcal{Q}^{(c)}$ via natural embeddings, or other similar quasi-orders for which the associated supply of isolated minimal antichains can be explicitly described?*

As we have mentioned there is a "doubling" construction which produces antichains by an even simpler process, which must be taken into account.

We are suggesting here that rather than just blindly following the "bootstrap" approach of the finiteness theorem, one can separate the issues out somewhat. Thus one can systematize more fully the search for appropriate antichains, and then return to the broader question of whether for $\mathcal{Q}$ as a whole, we are looking for a few families of appropriate constructions, or something considerably more chaotic.

Problem 8. *Are there encodings of $\mathcal{Q}^{(2)}$ or more generally $\mathcal{Q}^{(c)}$ into the quasiorder $\mathcal{P}$ of finite permutations which convert the almost periodic antichains into isolated minimal antichains of $\mathcal{P}$?*

One can ask more fundamental questions about these problems, for which the tools of logic and descriptive set theory are relevant. One way of posing our problem is in terms of the finiteness theorem; we ask whether the antichains involved can be described effectively, and uniformly. Results of logic suggest that these problems are highly nontrivial in general, even if the quasiorders involved are effectively given, and wellfounded (in fact strongly wellfounded: each element has only finitely many predecessors, as is the case when dealing with finite structures).

Harvey Friedman has shown the following (private communication). We use the term "locally finite" for a quasiorder such that each element lies above a finite set of elements, up to equivalence. This is a considerable strengthening of well-foundedness.

Fact 4.1 (Friedman)**.** *There is an effectively given (elementary recursive) and locally finite partial order $\mathcal{Q}$ such that the set $WQO_{\mathcal{Q},1}$ of constraints $c \in \mathcal{Q}$ for which the ideal $\mathcal{Q}_c$ is wellfounded is a complete Π^1_1 set.*

Here we invoke the definability hierarchy of logic, which is a good way of classifying sets which are radically non-computable. A set of integers is Π^1_1 if it has a definition beginning with one universal quantifier over sets of integers, with any remaining quantifiers being over integers. Here the natural definition of the set $WQO_{\mathcal{Q},1}$ has this form: in saying there is no infinite antichain, one quantifies over arbitrary subsets of $\mathcal{Q}$. A complete Π^1_1 set is of maximal complexity among all Π^1_1-definable sets (with respect to algorithmic reductions, specifically *many-one* reductions). For our purposes, the main point would be that one cannot effectively determine which constraints c

satisfy our condition (belong to $WQO_{\mathcal{Q},1}$). Indeed, sets of integers which can be recognized by an algorithm lie close to the bottom of the hierarchy of logical complexity, as far as definability is concerned, being definable using little more than a single existential numerical quantifier.

Thus Fact 4.1 is a very sharp way of saying that the set $WQO_{\mathcal{Q},1}$ is undecidable—and that the set is intrinsically as complex as a set with such a definition can be. This gives us a model for the "bad" case of Decision Problem $A_{\mathcal{Q}}$, and raises more pointedly the question as to whether our *natural cases* could conceivably include cases just as bad. But to us this still seems highly unlikely.

Friedman has also announced the following in a private communication:

- there is a finite signature consisting of just constant and function symbols for which the same occurs.

- The Finiteness Theorem of §2, for the case $k = 1$, is equivalent, over a weak base theory (RCA_0), to the Π^1_1 Comprehension Axiom $\Pi^1_1 - CA0$.

The first of the these results comes close to the combinatorial context considered here, though there are differences between functional languages and relational languages in this context.

The second result corresponds to the fact that there is a direct proof of the Finiteness Theorem for the case of single constraints, for which it is sufficient to take as the starting point the set of all constraints for which the wqo property fails; and it makes rigorous the claim that one cannot carry out this argument with less at one's disposal.

We give the proof of Fact 4.1.

Lemma 4.2. *Let $\leq$ be the usual ordering of $\mathbb{N}$, and $\leq_1$ any linear ordering of $\mathbb{N}$. Define a partial order $n \leq_2 m$ by ($n \leq m$ and $n \leq_1 m$). Then $(\mathbb{N}, \leq_2)$ is a locally finite quasiorder with the following property: $\leq_2$ is a wqo of the ideal $\mathbb{N}_n$ if and only if $\leq_1$ is a well order of the initial segment $(i \in \mathbb{N} : i <_1 n)$.*

Here we write $\mathbb{N}_n$ for $(\mathbb{N}, \leq_2)_n$, which by definition is the ideal

$$\{m \in \mathbb{N} : n \not\leq_2 m\}$$

of $(\mathbb{N}, \leq_2)$.

Proof. Suppose the initial segment $(i \in \mathbb{N} : i <_1 n)$ is wqo and I is an infinite antichain in $\mathbb{N}_n$. Removing finitely many elements of I, we may suppose that all are larger than n in the natural order. Since $I \subseteq \mathbb{N}_n$, they are then smaller than n in the order $\leq_1$. Taking $i \in I$ to be minimal in the order $\leq_1$, and $j \in I$ larger in the natural order, we find $i \leq_2 j$, a contradiction.

Now suppose that the initial segment $(i \in \mathbb{N} : i \leq_1 n)$ is not well ordered and that $x_1, x_2, \ldots$ is a decreasing sequence with respect to $\leq_1$. Then for any n there is some $m > n$ such that $x_m > x_1, \ldots, x_n$ in the natural order, and thus x_m is incomparable with $x_1, \ldots, x_n$ in $(\mathbb{N}, \leq_2)$. It follows that the sequence contains an infinite antichain in $\mathbb{N}_n$. $\square$

Now for the partial order $(\mathbb{N}, \leq_2)$, the set $\{n \in \mathbb{N} : \mathbb{N}_n \text{ is not wqo}\}$ coincides with the set

$$\{n \in \mathbb{N} : \text{The initial segment below } n \text{ relative to } \leq_1 \text{ is not well ordered}\}$$

This is a typical representation of a complete Π^1_1 set in recursion theory, with $\leq_1$ recursive [47, Chapter 1]. Fact 4.1 follows.

4.2. Universality Problems

The universality problem for classes determined by a single connected constraint has become increasingly amenable to analysis, notably with the advent of the pruning lemma.

In the bounded degree case to say that one vertex is immediately algebraic over another simply means that they are adjacent. But in general immediate algebraicity is witnessed by additional vertices, and then as one attempts to analyze further the interaction of these witnesses requires analysis. For an example of the complications ensuing, consider the following instance, where the analysis is still not complete.

Problem 9 (Hairy Ball Problem). *We consider graphs C consisting of a complete graph K on n vertices v_i, with at most one path P_i adjoined to each vertex v_i. We ask whether* acl_C *is locally finite.*

We tend to think acl is indeed locally finite in all such cases, so that there is a corresponding universal graph. But if this is not the case, then with a few counterexamples in hand, we would be able to use the pruning method to reduce the general problem substantially.

We have already carried out the translation of this problem into a completely explicit one, so we repeat that here.

Problem 10. *[Hairy Ball Problem, Explicit Form] With C as specified, let G be a graph containing a sequence of vertices $(v_i)_{i\in\mathbb{Z}}$ such that the v_i lie along a 2-way infinite path Q which is free in G over the $(v_i)_{i\in\mathbb{Z}}$ in the sense of Definition 3.9 and such that each vertex v_i belongs to a graph C_i so that the pair (v_i, C_i) is isomorphic to a pair (v, C_v) in the notation above. Does it follow that C embeds into G?*

We now turn to a different topic. We took note in §3.6 of a reduction of universality problems in general to the case of graphs with a coloring of the vertex by two colors. So we ask the next question.

Problem 11 (Graph Reduction Problem)**.** *Is there a reduction of the problem of the existence of countable universal vertex colored graphs (for the case of two colors) to the same problem for graphs, where a finite set of forbidden substructures is allowed in each case?*

Here a reduction would be, in general, any algorithmic reduction, but what is envisioned in particular would be an interpretation of the broader class in the narrower, which is how such results are usually obtained. Such a result would complete the reduction of all problems of this type to the category of graphs, which is encouraging if one thinks the latter problem might be solvable, and discouraging if one thinks the former problem is not likely to be solvable. The most intriguing possibility is that the universality problem is solvable for graphs but not for general structures; but such a reduction would close the door on this possibility.

There is one more problem that we consider very attractive, which as far as we know has received no systematic attention.

Problem 12 (Permutation Patterns)**.** *Identify those permutation pattern classes for which there is a unique existentially complete countable permutation, up to isomorphism.*

A variety of decision problems in the context of permutation pattern classes are discussed in [46]. When permutations are viewed as structures, pattern classes are precisely the classes defined by forbidden substructures which are themselves permutations [7]. We may extend a pattern class by considering all the countable permutations (i.e., reorderings of an arbitrary countable ordered set) which obey the same constraints. Cameron has given the complete classification of the universal homogeneous permutations; these all have a very simple structures.

In this setting, a permutation P is existentially complete (relative to the given class) if for every finite subpermutation P_0 of P and every extension of P_0 to a finite permutation Q_0, if Q_0 embeds over P_0 into an extension of P in the class, then Q_0 embeds into P over P_0. Every countable permutation in a pattern class extends to a countable existentially complete permutation for that class, and thus there will be a universal permutation if and only if there is a universal existentially complete permutation. In particular, when there is a unique existentially complete permutation in the class, there is a universal one, and furthermore we have a canonical infinite permutation representing the pattern class. If we take our experience with graphs as a guide, we may expect that universal permutations arise mainly in this fashion—but it is an uncertain guide, as mentioned before. We alluded in §3.6 to the need for some foundational work on the model theory side.

One can see this topic touched on in the permutation class literature in a very direct sort of way (looking for infinite limits as permutations of $\mathbb{Z}$ or $\mathbb{Q}$). But adopting Cameron's strategy of viewing permutations as structures, and then applying the standard apparatus of model theory, one should be able to make something more comprehensive out of this idea.

4.3. Notes

Here we address some points that might strike a close reader as calling for further comment.

1. By a "combinatorial structure" we have in mind, roughly speaking, a structure in a finite relational language. However we also allow the imposition of symmetry conditions: for each complete quantifier-free type in the language, we allow the specification of a symmetry group. And it is reasonable to require that a relation never hold of an n-tuple whose entries are not distinct, as one may add a relation in fewer variables to cover that case.

The article [11] was written throughout in the language of graph theory, but goes over without significant change to structures in finite relational languages with symmetry conditions on the relations.

In combinatorial model theory it is often appropriate to allow finitely many functions with a uniform finiteness condition (with vector spaces over finite fields a typical instance), but we have not looked in this direction.

2. If A, B are combinatorial structures of the same type, with the domain of A contained in the domain of B, we call A a *substructure* of B if the relations on A are contained in the corresponding relations on B, and an *induced substructure* if the relations on A are the restrictions to A of the

corresponding relations on B. When symmetry conditions are imposed on the basic relations of a structure (as in Note 1) then the relations of A as well as B must satisfy those symmetry conditions. (Thus for example if we consider a structure in which a certain binary relation R is symmetric, then any induced substructure will be symmetric, but this is not necessarily the case for substructures, unless we choose to work in the category of symmetric structures.)

Our terminology follows the usage of graph theory: "subgraph" and "induced subgraph" are, respectively, "substructure" and "induced substructure" when graphs are encoded by symmetric binary relations, and the symmetry is included in the specification of the language.

But this terminology conflicts with the usage of model theory, where "substructure" corresponds to our "induced substructure" and there is no common term corresponding to our "substructure," though "weak substructure" would be natural.

We gave some substantial reasons to work with forbidden substructures rather than forbidden induced substructures in §3.6. On the one hand our general theory holds in the former case and definitely not in the latter; on the other hand we have relatively straightforward encodings of undecidable problems in the latter setting, and nothing similar (so far) in the former setting.

The formal setting for wqo problems is flexible: any interesting class of finite combinatorial objects with a natural well-founded quasi-order will do. Thus the fact that the class $\mathcal{Q}$ of tournaments is not closed under substructure is not of particular importance in the context of §2.

But universality problems are more sensitively related to the initial choice of forbidden substructures, and other issues (the latter illustrated by Problem 12).

3. In the treatment of $\mathcal{Q}^{(c)}$ we work with oriented paths. One can work with symmetric paths without much alteration, evidently, and when encoding antichains into binary relational structures with symmetric relations, one would presumably do so.

4. In subsection 2.3 we saw that for the wqo problem, the case of graphs and the case of tournaments are very different; as technically speaking tournaments are not actually an instance of "finite relational systems" in the sense of model theory, and graphs are, one might wonder whether the latter case is more typical. However, what is actually important here is the fact that in the language of tournaments (or digraphs, if one prefers) there are

two positive relations on pairs: $E(x, y)$ and $E(y, x)$. So for example if one considers graphs with two colors of edges, encoded by relations $E_1(x, y)$ and $E_2(x, y)$, this should be similar to the case of tournaments.

Recall that the *Gaifman graph* (M, E) associated to any relational system $(M, \dots)$ has M as its set of vertices, with two vertices are adjacent if and only if they are part of a tuple related by at least one relation in M. One way to get antichains of structures is by ensuring that the corresponding Gaifman graphs are antichains; but on the other hand one also has antichains of structures whose Gaifman graphs are paths, and in the case of tournaments all the associated Gaifman graphs are complete.

5. Our "default" notion of universality is the strong one, where we require universal graphs to admit embeddings of other graphs as induced subgraphs. We have also taken note of the weak notion of universality, in which the embedding must be as simply as a subgraph.

These two notions can diverge considerably—in Rado's context, if all one wants is weak universality, a complete graph would suffice—but for the question of their existence, as opposed to the details of their structure, the corresponding dividing line does not seem to change much. In fact the work to date has always been done in the strongest possible form: in positive cases *strongly* universal graphs are *constructed*, and in negative cases the existence of *weakly* universal graphs is *refuted*. The latter may take a little extra work in some cases, requiring some decoration of the algebraic closure of a finite set so as to guarantee incompatibility of the structures involved, in a suitable sense. The work in the case of trees in particular [10] would be a little simpler without this fillip, but the structure of the argument would be unchanged.

The case of bounded degree (3.1) is an exception to this pattern. Here one comes down almost immediately to the case in which the algebraic closure operator is not locally finite, while in other cases that first phase of the analysis is the main one.

6. We have restricted our attention to three notions of tameness: wqo, the existence of a universal graph, and a local finiteness condition equivalent to the existence of a canonical universal graph. In model theoretic terms, this last condition is the $\aleph_0$-categoricity of the model companion of the theory of the class. Any of the tameness criteria of model theory (notably, stability and its variations) give rise to analogous classification and decidability problems. In general, the first order theories which axiomatize constraints given by finite sets of forbidden substructures constitute a class of universal theories that lend themselves to particularly systematic analysis. More precisely,

one considers the class of model complete theories whose universal part expresses a set of constraints given by finitely many forbidden substructures. Within this latter class, one expects the study of any natural model theoretic property to lead back to purely combinatorial problems involving the set of finite forbidden substructures from which the theory is derived. Experience in model theory would suggest that stability, in particular (or perhaps more the broader notion of simplicity), should have connections with more concrete combinatorial issues.

But we have not pursued this line of thought.

5. Appendix 1: Universality with Bounded Degree

In this Appendix we will elaborate on our previously sketched proof of the following.

Proposition (3.4)**.** *For constraint sets $\mathcal{C}$ including some star, the problem of the existence of a weakly universal countable $\mathcal{C}$-free graph is decidable. A weakly universal $\mathcal{C}$-free graph will exist if and only if one of the following conditions holds:*

1. *$\mathcal{C}$ contains a path;*
2. *$\mathcal{C}$ contains generalized 3-star $S(k_1, k_2, k_3)$ consisting of a central vertex v_0 and paths P_i of length k_i for $i = 1, 2, 3$ attached to v_0. In addition, any maximal infinite connected $\mathcal{C}$-free graph is almost periodic.*

We will concentrate on the structural analysis involved, and comment afterward on the issues involved in making this more constructive.

5.1. Block structure of $\mathcal{C}$-free graphs

We first go over some standard graph theory used in the structural analysis.

Definition 5.1.

1. A *cut vertex* of a connected graph is a vertex whose removal disconnects the graph.
2. A graph is 2-connected if it is connected and has no cut vertices.

3. An induced subgraph of a graph is a *block* if it is a maximal 2-connected induced subgraph.

4. The *tree of blocks* associated to a connected graph G has as vertices the blocks and the cut points of G, with the incidence relation as edge relation.

To justify the terminology one checks that the graph structure on the tree of blocks T is in fact a tree. This holds because the union of the blocks lying in any cycle of T would itself be 2-connected, contradicting the maximality of the blocks.

Lemma 5.2. *Let G be connected graph of infinite diameter omitting some generalized 3-star $S = S(k_1, k_2, k_3)$. Let $k = \max(k_1, k_2, k_3)$. Then the diameter of any block of G is at most k.*

We remark that this does not work with G of finite diameter, e.g. a large cycle.

Proof. Let B be a block of G, and suppose $a, b \in B$ lie at distance greater than k. There are two disjoint paths P_1, P_2 joining a to b in B and each has length at least k. Let C be the cycle formed by $P_1 \cup P_2$. Let δ be the diameter of C.

Take a vertex $c \in G$ with $d(a, c) > \delta + k$. Let Q be a path of minimal length connecting c to C. Then the length of Q is at least k. Then $Q \cup C$ contains S and we have a contradiction. $\square$

Corollary 5.3. *Let G be a connected infinite graph omitting some star and some generalized 3-star. Then the blocks of G have bounded order.*

Lemma 5.4. *Let G be a connected infinite graph omitting some star and some generalized 3-star S. Then there is a sequence B_i of blocks of G indexed by $I = \mathbb{N}$ or $\mathbb{Z}$, so that B_i meets exactly $B_{i\pm1}$ (with the obvious exception for $i = 0$, $I = \mathbb{N}$), and so that the connected components of $G \setminus \bigcup_i B_i$ are of bounded order.*

Proof. Let T be the tree structure induced on the blocks of G: its vertices are the blocks and cut points of G, with incidence as the edge relation. This tree omits the generalized 3-star S and is itself infinite. We claim that T contains a path P of order type $\mathbb{N}$ or $\mathbb{Z}$ such that $T \setminus P$ decomposes into trees of bounded order.

Let $S = S(k_1, k_2, k_3)$ and $k = \max(k_1, k_2, k_3)$. For any vertex v of T, the connected components of $T \setminus \{v\}$ are trees of bounded degree, and all but at most two contain no path of length $2k$; otherwise we find three disjoint paths of length k connecting to v.

Let P be the set of vertices of T for which there are, in fact, two components of $T \setminus \{v\}$ containing a path of length $2k$. T induces a tree structure on P, with edges in P corresponding to paths in T. If a vertex v in P has at least three neighbors v_1, v_2, v_3 in P, then there are components T_i of $T \setminus \{v_i\}$ such that T_i contains a path of length $2k$ and does not contain v. Then T_1, T_2, T_3 are contained in distinct connected components of $T \setminus \{v\}$ and we contradict the definition of P.

On the other hand, if $u, v \in P$ then the path connecting them lies in P as well.

So P is a path.

Consider the connected components of $T \setminus P$. If one of them is sufficiently large, it will contain a path of length $4k + 1$. Then the midpoint will belong to P, a contradiction. So $T \setminus P$ breaks up into trees of bounded order. In particular P is infinite, and can be indexed by $\mathbb{N}$ or by $\mathbb{Z}$.

For any vertex of $T \setminus P$ that lies between two vertices v_1, v_2 of P, as above we find components T_1, T_2 of $T \setminus \{v\}$ containing v_1, v_2 respectively together with a path of length $2k$. So P is a convex subset of T. So P represents a sequence of blocks and cut points, each incident with the next, and pulling back from T to G gives the claim. □

Apart from the choice of a numerical parameter, the path P constructed in the previous proof was obtained canonically from the graph G.

Notation 5.5. *With the hypotheses of Lemma 5.4, and with k fixed, let P be the path in the tree of blocks constructed in the proof. Let (P_B, P_V) be the partition of the vertices of P into blocks and cut vertices. Each of P_B and P_V may be construed as a path in which the remaining vertices of P encode edges (with the exception of the endpoint of P if there is one).*

At this point we can see the encoding of G by an infinite word in a fixed finite alphabet, indexed by the vertices of P. The next lemma is mainly a matter of establishing notation.

Lemma 5.6. *Let G be a graph, P an infinite path contained in the tree of blocks T associated with G, and P_V the set of cut vertices in P. For C a* *connected component of $T \setminus P_V$, let $\hat{C}$ be the union of the blocks in C. Then we have the following.*

1. *$\hat{C}$ meets P_V in one vertex or in two adjacent vertices of P_V.*

2. *As C varies over the connected components of $T \setminus P_V$, $\hat{C} \setminus P_V$ varies over the connected components of $G \setminus P_V$, with each such component occurring once.*

Proof. The first statement is a transparent statement about trees. A connected component C of $T \setminus P_V$ will have vertices adjacent to one or more vertices of P_V. Either C will contain a vertex on P and the two adjacent vertices on P_V will be the ones in question, or C will contain a vertex of $T \setminus P$ adjacent to a vertex of P_V, and the latter is then unique. A vertex v of C adjacent to a vertex a of P_V represents a block containing a, so in this way $\hat{C}$ picks up the neighboring vertices of P_V.

For the second statement, it is evident that the various $\hat{C}$ are connected, and their union contains all blocks of G, hence all vertices and all edges of G. Furthermore the sets $\hat{C} \setminus P_V$ are pairwise disjoint. This proves the second point. □

Notation 5.7. *Assume the graph G satisfies the hypotheses of Lemma 5.4, and let T be the tree of blocks associated with G, with $P_V \subseteq P \subseteq T$ corresponding.*

1. *For a, b adjacent vertices of P_V, set*

$$\begin{aligned} G_a &= \bigcup_{\hat{C} \cap P_V = \{a\}} \hat{C} \\ G_{a,b} &= \bigcup_{\hat{C} \cap P_V = \{a,b\}} \hat{C} \end{aligned}$$

2. *Fix an orientation of P_V so that P_V is an ordered path of type $\mathbb{N}$ or $\mathbb{Z}$, and in particular carries a successor relation. For $a \in P_V$ with successor $b \in P_V$, set $\mathcal{G}_a = (G_a \cup G_{a,b}, a, b)$, a finite graph with two distinguished base points.*

3. *Let Σ be the alphabet consisting of the isomorphism types of structures (G_0, u_1, u_2) consisting of finite graphs of bounded order with two distinguished base points, where the bound is a function of k which bounds*

the sizes of all possible $\mathcal{G}_a$. Let $W = W_G$ be the infinite word $(\mathcal{G}_a)_{a \in P_V}$ in the alphabet Σ.

Observe that after appropriate numerical parameters have been fixed to make these constructions canonical, two connected graphs satisfying our hypotheses will be isomorphic if and only if they correspond to the same infinite word W, allowing of course for translation of the index set it if it is of type $\mathbb{Z}$.

5.2. Proposition 3.4

With this notation in hand we can return to the proof of Proposition 3.4.

Proof. Let G be an infinite maximal connected $\mathcal{C}$-free graph, and $W = W_G$ the associated infinite word describing the construction of G relative to a suitable sequence of cut vertices. We must show that if G is not almost periodic, then there are $2^{\aleph_0}$ such graphs G. We are interested in finite segments of W, that is finite contiguous subwords.

At this point another numerical parameter comes into play: a bound K on the size of all constraints in $\mathcal{C}$. We need to see how the properties of G are reflected by the associated word W, and we claim that this is *local*, that is, depends only on the local structure of W, namely its subwords of length at most K.

Any connected subgraph of G of order, or for that matter diameter, at most K lies in a portion of G encoded by a segment of W of length at most K. Thus the $\mathcal{C}$-freeness is local, involving only the set of such segments occurring in W. Maximality is for the most part a similar condition. An edge not in G would either involve two points at a large distance, in which case the generalized 3-star occurring as a constraint already rules out the existence of such an edge, or else would involve two vertices v, v' of G at a uniformly bounded distance. In this case, maximality of G requires that insertion of the edge (v, v') would result in an embedding of some graph in $\mathcal{C}$ into the extended graph, and as the constraints are connected graphs, these would involve small graphs connected to v and v', and hence again encoded into short segments of W, containing suitable segments around v and v', and the segment from v to v'.

So for any infinite word W' whose segments of appropriately bounded length occur as segments of W, we may construct a graph G' with associated word W', which will then be an infinite maximal connected $\mathcal{C}$-free graph.

Now there is at least one suitably long finite segment w which occurs infinitely often in W. In particular there is a segment $ww'w$ of W with both w and w' suitably long. Then any short subword of $(ww')^\infty$ will be a subword of ww' or $w'w$, and hence of W.

We claim that either the word W is itself almost periodic with period ww', or that we can find an uncountable set of distinct words with the same local structure.

Suppose that W is not almost periodic. Then there is a finite segment ww'' with w'' at least as long as w', such that ww'' does not embed into $(ww')^\infty$. (Begin with any specific occurrence of w in W.) There is then a power $(ww')^n$ of ww' which does not embed in $(ww'')^\infty$. Consider the words $(ww')^n$ and ww'', which we label α, β, choosing the notation so that the word α is no longer than the word β. We record the relevant properties.

1. α, β are finite words, with α no longer than β.
2. α does not embed in β^∞.
3. Any 2-way infinite product $\prod_i w_i$ with $w_i \in \{\alpha, \beta\}$ encodes an infinite maximal connected $\mathcal{C}$-free graph.

Now consider products of words $w_i = \alpha\beta^{n_i}$ ($i \in \mathbb{Z}$, $n_i \geq 2$, even). Two such products W_1, W_2 corresponding to sequences n_i, n_i' will be translates of each other if and only if the sequences (n_i), (n_i') are themselves translates of each other. Indeed, let f embed W_1 into W_2. Then the image of each segment of type α in one of the words w_i has to meet some segment of type α in one of the words w_j, and while in principle these two segments may overlap properly, as α is no longer than β this allows at worst $|n_i - n_j'| \leq 1$, so by taking the entries even we ensure that they match, after which it becomes clear that the map from i to j determined by f is a translation.

So we arrive at $2^{\aleph_0}$ distinct words from which $2^{\aleph_0}$ nonisomorphic maximal connected $\mathcal{C}$-free graphs can be constructed, and there is no universal $\mathcal{C}$-free graph in this case. □

For the issue of decidability, one needs mainly in the concluding phase of the argument to determine how far from strict periodicity the sequence needs to be to produce the final construction of $2^{\aleph_0}$ graphs. If one is not much concerned about complexity issues it suffices to argue that there is such a bound, and also that once such a bound can be reached, it can be recognized. One can certainly enumerate the alphabet Σ (on the first pass, it is enough to give a finite alphabet containing Σ) and a list of all the words in Σ up to some finite bound which correspond to $\mathcal{C}$-free graphs, and which also satisfy an appropriate weak maximality condition, with respect to vertices lying well away from the boundary of the word. After that, one needs only to determine whether there is an infinite product of these words which is not almost periodic.

6. Appendix 2: Unarity vs. Universality

This appendix is devoted to a proof of the following.

Theorem 3. *There is a finite set $\mathcal{C}$ of finite connected graphs such that*

1. *The associated operation* $\mathrm{acl}_{\mathcal{C}}$ *of algebraic closure (in $\mathcal{C}$-free graphs) is locally finite;*
2. *In particular, the class $\mathcal{A}_{\mathcal{C}}$ of countable $\mathcal{C}$-free graphs has a universal member (Theorem 2); but—*
3. *The operation of algebraic closure* $\mathrm{acl}_{\mathcal{C}}$ *associated with $\mathcal{C}$ is not unary.*

The fact that we can have a universal $\mathcal{C}$-free graph without unarity undoubtedly complicates the issue of determining systematically when a universal countable $\mathcal{C}$-free graph exists. This phenomenon seems unlikely in the case of 1 constraint, but is not yet ruled out.

The idea of the proof is simple enough: consider the cartesian power X^2 as a structure $A = (X^2 \cup X; \pi_1, \pi_2)$ with π_1, π_2 the projection maps. If $S \subseteq A$ then taking $X_S = (S \cap X) \cup \pi_1(S) \cup \pi_2(S)$, we have $\mathrm{acl}(S) = X_S \cup X_S^2$. So this operation is locally finite and not unary, and if we can encode this situation faithfully into a graph structure, these properties will be inherited. It will then suffice to extract a finite number of forbidden substructures which are sufficient to carry through the same analysis.

6.1. Graph structures on $[X]^2$

Since we are working with graphs, we are forced to allow some graph structure on X^2, so we begin by considering a class of structures that encodes such a structure, except that we will replace the set of ordered pairs of X^2 by the set of unordered pairs, denoted $[X]^2$.

Definition 6.1. Let $\mathcal{A}$ be the class of countable structures of the form

$$A = (X; E)$$

where $E = E(x_1, x_2, x_3, x_4)$ is a 4-place relation on X which encodes a triangle-free graph on $[X]^2$ (in particular, E is invariant under permutations of the variables in the Klein 4-group), and such that $E(x_1, x_2, x_3, x_4)$ never holds if two of the entries are equal.

The next result is analogous to the existence of a universal triangle-free graph.

Lemma 6.2. *There is a universal homogeneous structure in $\mathcal{A}$.*

Such results are typically proved using Fraïssé's theory. For simplicity we formulate that theory in a very special case, adequate for our present purpose, as follows.

Fact 6.3. *Let $\mathcal{F}$ be a class of finite structures having finitely many relations and no functions, and assume that the following hold.*

- *If A is in $\mathcal{F}$, then any structure isomorphic to an induced substructure of A is also in $\mathcal{F}$.*
- *The disjoint union of two structures in $\mathcal{F}$ is again in $\mathcal{F}$.*
- *If $A_0, A_1, A_2 \in \mathcal{F}$ and $A_0 \subseteq A_1, A_2$ then the disjoint union of A_1 with A_2 over A_0 (denoted $A_1 \cup_{A_0} A_2$) is again in $\mathcal{F}$.*

Let $\mathcal{A}$ be the class of countable structures whose induced substructures belong to $\mathcal{F}$. Then there is a universal structure in $\mathcal{A}$.

Our lemma then follows directly.

Proof. The necessary properties are all immediate. In the case of amalgamation, we may suppose that the structures A_1, A_2 are taken so that their intersection is A_0, that is the disjoint union over A_0 is just the union. Then the relation E on $A_1 \cup A_2$ is $E_1 \cup E_2$, and this represents a graph on $[A_1 \cup A_2]^2$ in which vertices $\{a, b\}$ for which $a \in A_1 \setminus A_0$, and $b \in A_2 \setminus A_0$ are isolated. $\square$

6.2. Encoding by Graphs

The next step is to pick a particular encoding of our exotic structures by ordinary (simple, symmetric) graphs.

Definition 6.4. Let $A = (X, E) \in \mathcal{A}$. Then Γ_A is the graph with vertex set $(X \times \mathbb{Z}/3\mathbb{Z}) \cup [X]^2$ and with edge relation the symmetric closure of the following relation $\sim$:

$$\begin{cases} (x,i) \sim (x,j) & x \in X,\ i,j \in \mathbb{Z}/3\mathbb{Z},\ i \neq j \\ (x,0) \sim \{x,y\} & x,y \in X \text{ distinct} \\ \{x_1,x_2\} \sim \{x_3,x_4\} & x_i \in X\ (i=1,2,3,4) \text{ and } E(x_1,x_2,x_3,x_4) \text{ holds} \end{cases}$$

Then all of the graphs Γ_A have the following five properties, with the first four properties corresponding to negative constraints.

1. Distinct triangles are disjoint.

2. No vertex v has three neighbors such that each one lies on a triangle.

3. In any triangle, at most one vertex has a neighbor not in the triangle.

4. For any two distinct vertices v_1, v_2, each of which lies on a triangle, there is at most one vertex adjacent to both.

5. Any vertex v has at least two neighbors, each of which lies on a triangle; and for any two distinct triangles, there is a vertex with a neighbor on each one.

We let $\mathcal{A}^*$ be the class of countable graphs satisfying conditions (1–4).

Remark 6.5. *There is a set $\mathcal{C}$ of finite connected graphs such that $\mathcal{A}^*$ coincides with the class $\mathcal{A}_{\mathcal{C}}$ of countable graphs omitting each member of $\mathcal{C}$ (as a subgraph).*

More specifically, condition (1) is expressed by two such constraints, and modulo condition (1), conditions (2,3,4) are expressed by one constraint each.

Lemma 6.6. *Let $\Gamma \in \mathcal{A}^*$. Then there is a structure $A \in \mathcal{A}$ such that Γ embeds into Γ_A as an induced subgraph.*

Proof. For any vertex $v \in \Gamma$, let Γ'_v be the graph obtained from Γ by adjoining three vertices (a, b, c) forming a triangle, and one further edge (a, v).

Claim: If in Γ v does not have two neighbors such that each one lies on a triangle, then $\Gamma'_v \in \mathcal{A}^*$. Observe that the hypothesis on v includes the assumption that v itself does not lie on a triangle in Γ. The claim then follows by inspection of the conditions (1-4) in Γ'_v.

Applying the claim iteratively, Γ embeds as an induced subgraph into a graph $\bar{\Gamma} \in \mathcal{A}^*$ such that every vertex $v \in \bar{\Gamma}$ has two neighbors such that each lies on a triangle.

Now in $\bar{\Gamma}$ we consider the set X of triangles contained in $\bar{\Gamma}$, and the set Y of vertices of $\bar{\Gamma}$ not lying on triangles. We define a structure $A = (X; E)$ with underlying set X and with $E(x_1, x_2, x_3, x_4)$ defined by:

$$\exists v_1, v_2 \in Y \, \exists u_i \in x_i \, (i = 1, 2, 3, 4) \, v_1 \sim u_1, u_2 \& v_2 \sim u_3, u_4 \& v_1 \sim v_2$$

By conditions (1–3) we have $A \in \mathcal{A}$.

Furthermore $\bar{\Gamma}$ embeds as an induced subgraph of Γ_A. For $y \in Y$, we map y to the unordered pair $\{T_1, T_2\}$, where T_1 and T_2 are the two triangles which contain neighbors of y. And for any triangle T of $\bar{\Gamma}$, we map it into the corresponding triangle $\{T\} \times \mathbb{Z}/3\mathbb{Z}$ of Γ_A, in such a way that if T has a vertex v with a neighbor off T, then v corresponds to $(T, 0)$. $\square$

6.3. Universality and Non-unarity

Lemma 6.7. *$\mathcal{A}^*$ contains a (strongly) universal structure.*

Proof. Let $A \in \mathcal{A}$ be universal. We claim that Γ_A is universal in $\mathcal{A}^*$.

Let $\Gamma \in \mathcal{A}^*$ and let $B \in \mathcal{A}$ be taken so that Γ is isomorphic with an induced subgraph of Γ_B. Take an embedding of B into A as an induced substructure. This gives an embedding of Γ_B into Γ_A as an induced substructure. $\square$

Now we claim that the operation $\mathrm{acl}_{\mathcal{C}}$ associated with $\mathcal{A}^*$ is not unary.

Lemma 6.8. *Let $\Gamma \in \mathcal{A}^*$ and suppose that Γ satisfies condition* (5)*. Then $\Gamma = \mathrm{acl}_{\mathcal{C}}(\Gamma)$, relative to any extension Γ' of Γ in $\mathcal{A}^*$ (with the Γ an induced substructure of Γ').*

Proof. It suffices to show that if Γ is an induced substructure of $\Gamma_1 \in \mathcal{A}*$, then the free amalgam $\Gamma^* = \Gamma_1 \sqcup_\Gamma \Gamma_1$ (the disjoint union over Γ) is in $\mathcal{A}*$. As a matter of notation, it will be convenient to write Γ_2 for the second copy of Γ_1, bearing in mind that $\Gamma_1 \cap \Gamma_2 = \Gamma$.

We claim first that any triangle T in Γ^* which meets Γ is contained in Γ. Let v be a vertex in $T \cap \Gamma$. Then v has two neighbors $v_1, v_2 \in \Gamma$ which lie on triangles of Γ. So by condition (2) T is contained in Γ.

In particular, condition (1) holds in Γ^*.

For condition (2), suppose that the vertex $v \in \Gamma^*$ has three neighbors v_1, v_2, v_3, each lying on a triangle. If $v \in \Gamma_1 \setminus \Gamma$ then $v_1, v_2, v_3 \in \Gamma_1$ and we violate condition (2) in Γ_1. So we may suppose $v \in \Gamma$. By our hypothesis on Γ, we may suppose $v_1, v_2 \in \Gamma$. Then the existence of v_3 violates condition (2) in Γ_1 or Γ_2.

Condition (3) for Γ^* is clear, but it is worth noticing that it depends on the fact that the two factors Γ_1, Γ_2 are isomorphic over Γ.

Turning to condition (4), suppose that $v_1, v_2 \in \Gamma^*$, each one lies on a triangle in Γ^*. If $v_1, v_2 \in \Gamma$ then there is a vertex $u \in \Gamma$ adjacent to both, and condition (4) for Γ^* follows from condition (4) for Γ_1, Γ_2. Suppose therefore that $v_1 \in \Gamma_1 \setminus \Gamma$. Then any neighbor of v_1 in Γ^* lies in Γ_1, and hence, by condition (5) in Γ, must lie in $\Gamma_1 \setminus \Gamma$. So if v_1, v_2 have a common neighbor u then both u and v_2 must lie in Γ_1, and hence u is unique by condition (4) in Γ_1. □

Proposition 6.9. *The algebraic closure relation* $\mathrm{acl}_{\mathcal{C}}$ *associated with the class* $\mathcal{A}^*$ *is not unary.*

Proof. Consider the graph on seven vertices consisting of two triangles T_1, T_2 together with a vertex v adjacent to a unique vertex on each.

Evidently $v \in \mathrm{acl}_{\mathcal{C}}(T_1 \cup T_2)$. On the other hand, the triangles T_1, T_2 are algebraically closed in view of the previous lemma. □

With this the proof of Theorem 3 is complete.

7. Appendix 3: Forbidden Induced Graphs

Our aim in the present section is to prove the following.

Theorem 4. *The universality problem for graphs with finitely many forbidden induced subgraphs is undecidable.*

Or equivalently: *the problem of determining whether a universal theory of graphs has a universal countable model is undecidable.*

Our method is to encode tiling problems, which involves covering the grid $\mathbb{Z}^2$ by a finite set of tile types, respecting some horizontal and vertical tiling constraints. As an intermediate step, we introduce *string tiling* problems, which are similar problems with the tile colors represented by bit strings (thereby reducing to the case of two tiles, with slightly more complicated tiling rules). It will be convenient to state a technical condition somewhat stronger than undecidability for both tiling problems, namely *inseparability* of the class of unsolvable tiling problems from the class of tiling problems with many solutions. All of this involves minor tweaks of the original tiling problem. Here one would expect "many" solutions to mean $2^{\aleph_0}$, but in fact we require a slightly sharper condition: $2^{\aleph_0}$ *incomparable* solutions, with respect to the pointwise comparison of colorings $f : \mathbb{Z}^2 \to \{0, 1\}$.

With that in place, we will introduce the class of graphs associated with a particular string tiling problem, in such a way that unsolvable cases of the tiling problem will give classes of graphs satisfying a finiteness condition that results in a universal structure, while the existence of many tilings refutes the existence of a universal structure.

The preparation with regard to tiling problems in §§7.1–7.2 is extremely natural, and can probably be taken in at a glance. The encoding by graphs was more elusive.

7.1. Tiling Problems

The intuition in the following definition is that we will have a finite set of colors C, an infinite collection of tiles of each color, and that we will be attempting to tile the plane $\mathbb{Z}^2$ completely with the specified tiles, while avoiding certain clashes of colors in the horizontal and vertical directions. There the structure on $\mathbb{Z}^2$ is a directed graph structure corresponding to the successor relation on $\mathbb{Z}$.

Definition 7.1.

1. A *tile set* is a structure $(C; H, V)$ with C finite and $H, V \subseteq C^2$.
2. Let $\mathcal{T} = (C; H, V)$ be a tile set. A function

$$f : \mathbb{Z}^2 \to C$$

is a $\mathcal{T}$*-tiling* if for all $m, n \in \mathbb{Z}$ we have

$H(f(m, n), f(m + 1, n))$ and
$V(f(m, n), f(m, n + 1))$

3. The *tiling problem* is the problem of deciding, for a given tile set $\mathcal{T}$, whether or not there is a $\mathcal{T}$-tiling of $\mathbb{Z}^2$.

See §3.6 for comments on the relationship of these tiles with Wang's dominoes.

Fact 7.2 ([2]). *The tiling problem is undecidable.*

As we said at the outset, this fact needs to be recast for our present purposes, using the following notion of algorithmic inseparability.

Definition 7.3. Let A, B be two disjoint sets whose elements are finite structures of some type, for example tile sets. We say that A and B are *algorithmically inseparable* if there is no computable function ϕ with domain all finite structures of the specified type, which separates A and B in the sense that $\phi = 0$ on A and $\phi = 1$ on B.

Lemma 7.4. *Let W_- be the set of all tile sets $\mathcal{T}$ for which there is no $\mathcal{T}$-tiling of $\mathbb{Z}^2$, and let W_+ be the set of all tile sets $\mathcal{T}$ for which there are $2^{\aleph_0}$ $\mathcal{T}$-tilings of $\mathbb{Z}^2$. Then W_+ and W_- are algorithmically inseparable.*

Proof. Short version: if you replace each tile in a set by two tiles, then every solution to the original tiling problem corresponds to $2^{\aleph_0}$ solutions to the new tiling problem.

More formally, if $\mathcal{T} = (C; H, V)$ is a tile set, define the *double* $\tilde{\mathcal{T}} = (\tilde{C}; \tilde{H}, \tilde{V})$ by

$$\begin{aligned} &\tilde{C} = C \times \{0, 1\} \\ &\pi : \tilde{C} \to C \text{ (projection)} \\ &\tilde{H} = \pi^{-1}[H],\, \tilde{V} = \pi^{-1}[V] \end{aligned}$$

(The map $\pi : \tilde{C}^2 \to C^2$ implicit in the last clause above is the map induced by π.)

Then $\tilde{\mathcal{T}}$-tilings $\tilde{T} : \mathbb{Z}^2 \to \tilde{C}$ project to $\mathcal{T}$-tilings via

$$\pi^*(\tilde{\mathcal{T}}) = \pi \circ \tilde{\mathcal{T}}$$

and $2^{\aleph_0}$ $\tilde{\mathcal{T}}$-tilings cover any single $\mathcal{T}$-tiling.

Thus if ϕ were a computable function separating W_+ and W_-, the function $\phi(\tilde{\mathcal{T}})$ would solve the Wang tiling problem. $\square$

7.2. String Tiling Problems

We will find it convenient to work with colorings of $\mathbb{Z}^2$ by two colors rather than by an arbitrary set of colors, so we introduce another set of tiling problems.

Definition 7.5.

1. A *string tile set* is a tile set $(C; H, V)$ in which C is a set of binary strings of fixed length: $C \subseteq 2^k$ for some k.

2. If $\mathcal{T} = (C; H, V)$ is a string tile set with $C \subseteq 2^k$, then a function
$$f : \mathbb{Z}^2 \to \{0, 1\}$$
is a $\mathcal{T}$*-coloring* if for all $m, n \in \mathbb{Z}$ the strings
$$\sigma_{m,n} = (f(m + i, n))_{i<k}$$
satisfy
$H(\sigma_{m,n}, \sigma_{m+k,n})$
$V(\sigma_{m,n}, \sigma_{m,n+1})$

3. If $f, g : \mathbb{Z}^2 \to \{0, 1\}$, we write $f \leq g$ if $f(m, n) \leq g(m, n)$ throughout $\mathbb{Z}^2$, and we refer to this as the pointwise partial order. In particular, f and g will be called *incomparable* if they are incomparable in this partial order.

We aim to convert the inseparability result for tiling problems to a very similar result for string tiling problems. Our encoding of tiling problems by string tiling problems is very direct, but as we have a number of technical constraints to observe, it takes some space to check thoroughly. This is the subject of the next lemma.

We first insert a word about *shifts* of tilings and colorings. For f a function with domain $\mathbb{Z}^2$ and $(m, n) \in \mathbb{Z}^2$, we define the shift f' of f by (m, n) via $f'(x) = f(x-(m, n))$. A shift of a $\mathcal{T}$-coloring f is again a $\mathcal{T}$-coloring, but if f is the bitstring representation of a tiling by tiles in 2^k, its shift by (m, n) will generally not be the bitstring representation of a tiling, unless m is a multiple of k. So we take this into account in our description of the correspondence between tiling and string tiling problems.

Lemma 7.6. *There is a transformation of tiling problems $\mathcal{T} = (C; H, V)$ into string tiling problems $\tilde{\mathcal{T}} = (\tilde{C}; \tilde{H}, \tilde{V})$ so that the following hold.*

1. *$\mathcal{T}$-tilings $T : \mathbb{Z}^2 \to C$ correspond naturally (and bijectively, up to shifts) to $\tilde{\mathcal{T}}$-colorings $f : \mathbb{Z}^2 \to \{0, 1\}$.*

2. *Any two $\tilde{\mathcal{T}}$-colorings $f, g : \mathbb{Z}^2 \to \{0, 1\}$ are incomparable in the pointwise partial order.*

Proof. We make a direct encoding of an arbitrary tile set $\mathcal{T}$ into a string tile set $\tilde{\mathcal{T}}$ so that $\mathcal{T}$-tilings and $\tilde{\mathcal{T}}$-colorings correspond, using strings of repeated 1's to mark tile boundaries, and using a set of incomparable strings to encode tile types. This will translate the inseparability result for tilings into the desired result. We now pass to the details.

Let $\mathcal{T} = (C; H, V)$ be a tile set. Take k so that $|C| \leq 2^k$ and assume that $C \subseteq 2^k$ is a set of bit strings. For $\sigma \in 2^k$ let $\bar{\sigma} \in 2^k$ be the complementary bit string, $\bar{\sigma}(i) = 1 - \sigma(i)$. Replacing each string $\sigma \in C$ by the string $\sigma\bar{\sigma}$, we may suppose that $C \subseteq 2^{2k}$ and the strings in C are incomparable.

Next we insert 0's to eliminate repeated occurrences of 1. So for $\sigma \in 2^{2k}$ we define $\sigma' \in 2^{4k+1}$ by

$$\sigma'(i) = \begin{cases} 0 & i \text{ even} \\ \sigma((i-1)/2) & i \text{ odd} \end{cases}$$

Let α_0 be the string 011, α_1 the string 110, and define

$$\tilde{\sigma} = \alpha_0 \sigma' \alpha_1$$

for $\sigma \in C$. We set $\tilde{C} = \{\tilde{\sigma} : \sigma \in C\}$ and $\tilde{k} = 4k + 7$.

Let $\tilde{H}$ be the set of all pairs $(\sigma, \tau) \in \tilde{C}^2$ which embed as a contiguous substring of some string $\tilde{\sigma}_1\tilde{\sigma}_2\tilde{\sigma}_3$ with σ_1, σ_2, σ_3 in C and $H(\sigma_1, \sigma_2)$, $H(\sigma_2, \sigma_3)$. Let $\tilde{V}$ be the set of all pairs σ, τ such that the $2 \times \tilde{k}$ array

$$\begin{pmatrix} \tau \\ \sigma \end{pmatrix}$$

embeds into an array of the form

$$\begin{pmatrix} \tilde{\tau}_1\tilde{\tau}_2 \\ \tilde{\sigma}_1\tilde{\sigma}_2 \end{pmatrix}$$

where the σ_i and τ_i are in C and we have $H(\sigma_1, \sigma_2)$, $H(\tau_1, \tau_2)$, $V(\sigma_i, \tau_i)$ $(i = 1, 2)$. Finally, set $\tilde{\mathcal{T}} = (\tilde{C}; \tilde{H}, \tilde{V})$.

To go from a $\mathcal{T}$-tiling $T : \mathbb{Z}^2 \to C$ to a $\tilde{\mathcal{T}}$-coloring $\tilde{T} : \mathbb{Z}^2 \to \{0, 1\}$, we replace the tiles by the corresponding strings, getting

$$\tilde{T}(m\tilde{k} + i, n) = \widetilde{T(m,n)}_i \ (i < \tilde{k})$$

where $\widetilde{T(m,n)}$ is the string encoding $T(m, n)$ and the subscript denotes its i-th entry.

By our construction, every $\mathcal{T}$-tiling produces a $\tilde{\mathcal{T}}$-coloring, and furthermore distinct $\mathcal{T}$-tilings give rise to incomparable $\tilde{\mathcal{T}}$-colorings. So we need only check that every $\tilde{\mathcal{T}}$-coloring arises in the intended manner, as the encoding of a $\mathcal{T}$-tiling.

Suppose therefore that $f : \mathbb{Z}^2 \to \{0, 1\}$ is a $\tilde{\mathcal{T}}$-coloring. We work mainly with the horizontal constraints $\tilde{H}$. We claim that for each n, the infinite string

$$\tau_n = (f(i, n))_{i \in \mathbb{Z}}$$

has a unique decomposition as a concatenation

$$\tau_n = \prod_{j \in \mathbb{Z}} \widetilde{\sigma^j} \quad (\sigma^j \in C)$$

This uniqueness should not be taken overly literally: any shift of the index set gives another parametrization of the same decomposition, a point which becomes more relevant when n varies. The horizontal constraints $\tilde{H}$ ensure that the pattern 11011 repeats regularly and that the strings occurring between successive instances encode elements of C, that is σ' for $\sigma \in C$. This gives us the desired unique decomposition.

Now we will decode the $\tilde{\mathcal{T}}$-coloring $f : \mathbb{Z}^2 \to \{0, 1\}$, making an appropriate shift at this point. Pick an occurrence of the pattern 110011 in τ_0, and shift the function f so that the second 0 occurs in position $(0, 0)$. Observe that the vertical constraint $\tilde{V}$ forces the occurrences of the pattern 110011 to be aligned vertically, and therefore after this shift the function f is the encoding of *some* function $T : \mathbb{Z}^2 \to C$. Then it is immediate that the constraints $\tilde{H}$ and $\tilde{V}$, and the distinguished positions of the patterns 110011, force T to be $\mathcal{T}$-tiling. $\square$

And we may now read off the inseparability result.

Lemma 7.7. *Let S_- be the set of string tile sets $\mathcal{T}$ for which there is no $\mathcal{T}$-coloring of $\mathbb{Z}^2$, and let S_+ be the set of string tile sets $\mathcal{T}$ with the following properties.*

(a) There are $2^{\aleph_0}$ $\mathcal{T}$-colorings of $\mathbb{Z}^2$;

(b) Any two $\mathcal{T}$-colorings f, g of $\mathbb{Z}^2$ are incomparable with respect to the pointwise partial order.

Then S_+ and S_- are algorithmically inseparable.

Proof. If the computable function ϕ separates S_+ and S_-, then using the function $\mathcal{T} \mapsto \tilde{\mathcal{T}}$ of Lemma 7.6, the computable function $\phi(\tilde{\mathcal{T}})$ separates W_+ and W_-. □

7.3. Encodings by Graphs

We now come back to universality problems for classes of graphs. This requires a transformation from string tile sets $\mathcal{T}$ to classes of countable graphs $\mathcal{A}_{\mathcal{T}}$. We begin by defining a particular encoding of colorings

$$f : \mathbb{Z}^2 \to \{0, 1\}$$

by graphs Γ_f.

Definition 7.8.

1. Let S be the symmetrized successor relation on $\mathbb{Z}$: $S(m, n)$ means $|m - n| = 1$.

2. Let $\tilde{\mathbb{Z}}$ be the graph obtained from $(\mathbb{Z}, S)$ by adjoining a cycle C_i of length $2(i \mod 6) + 3$ to the vertex $i \in \mathbb{Z}$.
   Explicitly, set $c_i = 2(i \mod 6) + 3$, represent the vertices of $\tilde{\mathbb{Z}}$ as pairs

   $$(i, v)\text{: } i \in \mathbb{Z},\ v \in \mathbb{Z}/c_i\mathbb{Z}$$

   and define the edge relation by: $E((i, v), (j, v'))$ iff

   $$\begin{aligned} j = i \pm 1,\ v = v' = 0 \\ \text{or } i = j, v = v' \pm 1 \end{aligned}$$

   In particular we identify $\mathbb{Z}$ with its image in $\tilde{\mathbb{Z}}$, writing i rather than $(i, 0)$.

3. Given $f : \mathbb{Z}^2 \to \{0,1\}$, let E_f be the extension of the edge relation on $\tilde{\mathbb{Z}}$ by the following relation on $\mathbb{Z}$

$$\{(a,b) \in \mathbb{Z}^2 : \exists x, y \in \mathbb{Z}\, f(x,y) = 1 \& \{a,b\} = \{6x, 6y+3\}\}$$

Let $\Gamma_f = (\tilde{\mathbb{Z}}, E_f)$

Remark 7.9. *For any $f : \mathbb{Z}^2 \to \{0,1\}$, the graph $E_f \restriction \mathbb{Z}$ is bipartite, with partition the even and odd integers. Thus the only odd cycles in Γ_f are the cycles $C_i = \{i\} \times \mathbb{Z}/c_i\mathbb{Z}$.*

Now we turn to the definition of the class $\mathcal{A}_{\mathcal{T}}$ of countable graphs associated with a string tile set $\mathcal{T}$. This class should contain all the graphs Γ_f for $f : \mathbb{Z}^2 \to \{0,1\}$ a $\mathcal{T}$-coloring, and should be defined by finitely many forbidden induced subgraphs. We need to collect some appropriate properties of the graphs Γ_f.

The first such constraints are as follows. Let $A_0 = \{c_i : 0 \leq i < 6\}$ be the set of sizes of odd cycles in the graphs Γ_f. Let R_0 be the cyclic successor relation on A_0 given by $R_0(c_i, c_{i+1 \bmod 6})$, R_1 the corresponding predecessor relation on A_0, R the symmetric relation $R_0 \cup R_1$, and R_2 the relation $R \cup \{(c_0, c_3), (c_3, c_0)\}$ Then the following constraints hold in the graphs Γ_f introduced above.

1. Any vertex belongs to at most one cycle whose order is in A_0.

2. For any cycle C whose order is in A_0, there is a most one vertex of C of degree greater than 2.

3. If C, C' are distinct cycles whose orders c, c' lie in A_0, and if some vertex of C is adjacent to a vertex of C', then $R_2(c, c')$ holds.

4. If C is a cycle of order $c \in A_0$, $v \in C$, $c' \in A_0$, and $R(c, c')$ holds, then there is at most one cycle of order c' containing a vertex adjacent to v.

Definition 7.10. Let Γ be any graph satisfying conditions (1–4).

1. Z_Γ is the set of vertices $v \in \Gamma$ of degree at least 3 which lie on some cycle C whose order is in A_0, Z'_Γ is the subset of Z_Γ of vertices lying on a triangle (recall $c_0 = 3$).

2. We define a successor relation S on Z_Γ as follows. Let E be the edge relation on Γ.

$$S = \{(a,b) \in Z_\Gamma^2 : E(a,b) \text{ and } a,b \text{ lie on cycles } C_1, C_2 \\ \text{of orders } c_1, c_2 \in A_0, \text{ with } R_0(c_1,c_2)\}$$

   The connected components of (Z_Γ, S) are oriented paths, or oriented cycles.

3. If $a, b \in Z'_\Gamma$ we define a partial function

$$\chi_{\Gamma,a,b} : \mathbb{Z}^2 \to \{0,1\}$$

   by

$$\begin{aligned}&\chi_{\Gamma,a,b}(m,n) = 1 \text{ iff}\\ &a' = S^{6m}a,\ b' = S^{6n+3}b \text{ are both defined, and } E(a',b')\end{aligned}$$

Observe that the domain of the function $\chi_{\Gamma,a,b}$ will be of the form $I \times J$ with $I, J \subseteq \mathbb{Z}$ intervals containing 0. Our last constraint is associated with a string tile set $\mathcal{C}$.

$(5)_\mathcal{T}$ The function $\chi_{\Gamma,a,b}$ is $\mathcal{T}$-admissible, for all $a, b \in Z'_\Gamma$.

By this we mean that the function satisfies the requirements on a $\mathcal{T}$-coloring, wherever it is defined.

Remark 7.11. *Constraints (1–4) correspond to finitely many forbidden subgraphs. Constraint* $(5)_\mathcal{T}$ *corresponds to finitely many induced subgraphs.*

The last point is perhaps unclear as m and n can be arbitrarily large. But since we can shift the base point (a,b) and the constraint $(5)_\mathcal{T}$ is a local one, the same condition can be expressed with bounded m and n.

Definition 7.12. Let $\mathcal{T}$ be a string tile set. Then $\mathcal{A}_\mathcal{T}$ is the set of countable graphs satisfying constraints (1–4) and $(5)_\mathcal{T}$.

7.4. Universality in $\mathcal{A}_\mathcal{T}$

Lemma 7.13. *Let $\mathcal{T}$ be a string tile set such that*

1. *There are $2^{\aleph_0}$ $\mathcal{T}$-colorings of $\mathbb{Z}^2$.*

2. *Any two distinct $\mathcal{T}$-colorings $f, g : \mathbb{Z}^2 \to \{0,1\}$ are incomparable in the pointwise partial order.*

Then there is no (weakly) universal graph in $\mathcal{A}_\mathcal{T}$.

Proof. Suppose $\Gamma \in \mathcal{A}_\mathcal{T}$ is weakly universal. For each $\mathcal{T}$-coloring $f : \mathbb{Z}^2 \to \{0,1\}$, fix an embedding

$$f^* : \Gamma_f \to \Gamma$$

as a subgraph.

As Γ is countable, there must be two distinct $\mathcal{T}$-colorings f_1, f_2 for which

$$f_1^*(i) = f_2^*(i) \text{ for } i = 0, 1 \in \mathbb{Z}$$

The constraints on Γ then force the embeddings f_1, f_2 to agree throughout $\mathbb{Z}$: $f_1^* \restriction \mathbb{Z} = f_2^* \restriction \mathbb{Z}$.

Let $\phi = f_1^* \restriction \mathbb{Z}$ and let $a = \phi(0)$. Then $f_1 \leq \chi_{\Gamma,a,a}$ pointwise. As $\chi_{\Gamma,a,a} : \mathbb{Z}^2 \to \{0,1\}$ is total and $\Gamma \in \mathcal{A}_\mathcal{T}$, χ_{Γ_a} is a $\mathcal{T}$-coloring and therefore by hypothesis $f_1 = \chi_{\Gamma,a,a}$. Similarly $f_2 = \chi_{\Gamma,a,a}$ and $f_1 = f_2$, a contradiction. $\square$

Now we consider tile sets $\mathcal{T}$ for which there is no $\mathcal{T}$-coloring, and we claim that in this case $\mathcal{A}_\mathcal{T}$ does contain a universal graph. We will need to apply the Fraïssé theory to a variation on the class $\mathcal{A}_\mathcal{T}$ equipped with additional functions.

Definition 7.14. Let $\mathcal{A}'_\mathcal{T}$ be the set of structures of the form

$$(\Gamma, E, f_1, f_2, f_3)$$

such that

1. $(\Gamma, E) \in \mathcal{A}_\mathcal{T}$

2. On Z_Γ, f_1 gives a partial successor function corresponding to the relation S, with $f_1(x) = x$ where the successor is not defined. Similarly f_2 represents the inverse of the successor function, with $f_2(x) = x$ where not otherwise defined.

3. For cycles C whose length is in A_0, f_3 gives a successor function on C (in other words, gives C an orientation); at vertices $v \in \Gamma$ not lying on such cycles, $f(v) = v$.

For $\Gamma \in \mathcal{A}'_{\mathcal{T}}$, and $A \subseteq \Gamma$, the substructure of Γ generated by A is the union of A, all connected components of (Z_Γ, S) which meet A, and all cycles of appropriate length meeting such a component.

Lemma 7.15. *Let $\mathcal{T}$ be a string tile set for which there is no $\mathcal{T}$-coloring of $\mathbb{Z}^2$. Then $\mathcal{A}'_{\mathcal{T}}$ has the following properties.*

1. *$\mathcal{A}'_{\mathcal{T}}$ is hereditary: closed under substructure and isomorphism.*

2. *The number of isomorphism types of finitely generated structures in $\mathcal{A}'_{\mathcal{T}}$ is countable.*

Proof. Condition (1) is clear.

For condition (2), since there is no $\mathcal{T}$-coloring of $\mathbb{Z}^2$, there is some maximal finite $L = L_{\mathcal{T}}$ for which there is a $\mathcal{T}$-coloring of $[0, L)^2$ (writing $[0, L)$ for $\{n \in \mathbb{Z} : 0 \leq n < L\}$). This bound follows by König's Tree Lemma, or a compactness argument.

Thus under our hypothesis on $\mathcal{T}$, the finitely generated structures in $\mathcal{A}'_{\mathcal{T}}$ are finite, and the claim follows. ☐

The foregoing lemma gives the more innocuous hypotheses of the Fraïssé theory. The next lemma addresses the main issue (amalgamation).

Lemma 7.16. *Let $\Gamma_0, \Gamma_1, \Gamma_2 \in \mathcal{A}'_{\mathcal{T}}$ with Γ_0 an induced substructure of Γ_1 and Γ_2. Then there is an* amalgam Γ *of Γ_1, Γ_2 over Γ_0 in $\mathcal{A}'_{\mathcal{T}}$. That is, we have embeddings $i_1 : \Gamma_1 \to \Gamma$ and $i_2 : \Gamma_2 \to \Gamma$ as induced substructures, agreeing on Γ_0.*

Proof. Form the disjoint union Γ° of Γ_1 and Γ_2 over Γ_0, with respect to vertices, edges, and the functions f_1, f_2, f_3. This is not yet in $\mathcal{A}'_{\mathcal{T}}$ for the following reason. If we take $a \in \Gamma_1 \setminus \Gamma_0$ and $b \in \Gamma_2 \setminus \Gamma_0$ and consider the function $\chi_{\Gamma^\circ,a,b}$, we find that this function is identically 0, which is not compatible with our constraints. We have to complete the definition of Γ by adding some edges encoding a suitable coloring of products $A \times B$ with A, B components of Z'_Γ with respect to the successor relation S^6 (and its inverse).

Note however that if A is a component of Γ_1 with respect to the successor function S^6 then for $a \in A$ the function $\chi_{\Gamma_1,a,a}$ gives a $\mathcal{T}$-admissible coloring of an interval of length $|A|$. This forces $|A| \leq L$. Thus we need only find encodings of $\mathcal{T}$-admissible colorings of products $A \times B$ with $|A|, |B| \leq L$, and by definition of L, there are such.

To make the last step a little more explicit: if $f : A \times B \to \{0, 1\}$ is a suitable coloring of $A \times B$, we add edges $(a, S^3 b)$ and $(S^3 b, a)$ that encode this (in particular we can encode $f(a, b) = 1$ and $f(b, a) = 0$ without conflict). $\square$

Corollary 7.17. *Let $\mathcal{T}$ be a string tile set. Then*

1. *If $\mathcal{T} \in S_+$ then there is no weakly universal graph in $\mathcal{A}_{\mathcal{T}}$.*
2. *If $\mathcal{T} \in S_-$, there is a strongly universal graph in $\mathcal{A}_{\mathcal{T}}$.*

Proof. We treated the first point in Lemma 7.13.

The second point requires some mopping up still. By Lemmas 7.15 and 7.16, the class $\mathcal{A}'_{\mathcal{T}}$ satisfies Fraïssé's conditions for the existence of a strongly universal (even universal homogeneous) structure; we have omitted explicit mention of the joint embedding property, but that is the case $\Gamma_0 = \emptyset$ in Lemma 7.16.

So let $(\Gamma^*, E, f_1, f_2, f_3)$ be universal and let $\Gamma' = (\Gamma^*, E)$ be the underlying graph. Then $\Gamma' \in \mathcal{A}_{\mathcal{T}}$, and we claim Γ' is universal. Taking $\Gamma \in \mathcal{A}_{\mathcal{T}}$, we expand it to a structure $\hat{\Gamma}$ in $\mathcal{A}'_{\mathcal{T}}$ by orienting the cycles whose lengths are in A_0 arbitrarily and then defining the functions f_1, f_2, f_3 correspondingly. Then $\hat{\Gamma}$ embeds into Γ^* as an induced substructure, and the restriction to Γ carries it into the graph Γ' as an induced substructure. $\square$

Now by Lemma 7.7 and Corollary 7.17, the problem of determining whether a given class of the form $\mathcal{A}_{\mathcal{T}}$ contains a universal structure is algorithmically undecidable. Thus Theorem 4 follows.

Acknowledgments

I thank Komjáth for introducing me to the universality problem in the form considered here, Harvey Friedman for drawing my attention to wqo, Kruskal's tree theorem, and its unusual strength from a logical perspective, and Vince Vatter for suggesting connections with permutations via Cameron's point of view in [7]. Debts to work by or with Latka, Shelah, Shi, and Tallgren are visible throughout the paper.

For comments made as I was writing the first draft of the present paper, I thank Sam Buss, and again Harvey Friedman. For useful comments on that

first draft I thank two energetic anonymous referees, whose remarks led to a substantial reworking of the presentation.

References

[1] M. Atkinson, M. Murphy, and N. Ruškuc, "Partially well-ordered closed sets of permutations," *Order* **19** (2002), 101–113.

[2] R. Berger, "The undecidability of the domino problem," *Mem. Amer. Math. Soc.* **66**, 72 pp., 1966.

[3] M. Bóna, *Combinatorics of permutations. With a foreword by Richard Stanley*, Discrete Mathematics and its Applications (Boca Raton). Chapman & Hall/CRC, Boca Raton, FL, 2004. xiv+383 pp.

[4] M. Bóna, *Editor's preface [Special issue on pattern avoiding permutations], Adv. in Appl. Math.* **36** no. 2 (2006).

[5] R. Brignall, "Antichains of Tournaments," slides from a talk, June 17, 2008, International Conference on Permutation Patterns, 2008; http://users.mct.open.ac.uk/rb8599/talks/PP08Talk.pdf

[6] R. Brignall, N. Ruškuc, and V. Vatter, "Simple permutations: decidability and unavoidable substructures," *Theoret. Comput. Sci.* **391** (2008), 150–163,

[7] P. Cameron, "Homogeneous permutations," in Permutation patterns (Otago, 2003), *Electron. J. Combin.* **9** (2002/03), Research paper 2, 9 pp.

[8] G. Cherlin, "The classification of countable homogeneous directed graphs and countable homogeneous n-tournaments," *Mem. Amer. Math. Soc.* **131** (1998), no. 621, xiv+161 pp.

[9] G. Cherlin and B. Latka, "Minimal antichains in well-founded quasi-orders with an application to tournaments," *J. Combin. Theory Ser. B* **80** (2000), 258–276.

[10] G. Cherlin and S. Shelah, "The tree conjecture for universal graphs," *J. Comb. Theory Ser. B* **97** (2007), 293–333.

[11] G. Cherlin, S. Shelah, and N. Shi, "Universal graphs with forbidden subgraphs and algebraic closure," *Advances in Applied Mathematics* **22** (1999), 454–491.

[12] G. Cherlin and N. Shi, "Graphs omitting a finite set of cycles," *J. Graph Theory* **21** (1996), 351–355.

[13] G. Cherlin and N. Shi, "Forbidden subgraphs and forbidden substructures," *J. Symbolic Logic* **66** (2001), 1342–1352.

[14] G. Cherlin and L. Tallgren, "Universal graphs with a forbidden near-path or 2-bouquet," *J. Graph Theory* **56** (2007), 41–63.

[15] J. Daligault, M. Rao, and S. Thomassé, "Well-quasi-order of relabel functions," *Order* **27** (2010), 301–315, Special Issue: ROGICS'08—International Conference on Relations, Orders and Graphs—Mahdia, Tunisia, May 2008 (eds., Nejib Zagula and Dwight Duffus).

[16] G. Ding, "Subgraphs and well-quasi-ordering," *J. Graph Theory* **16** (1992), 489–502.

[17] Z. Füredi and P. Komjáth, "On the existence of countable universal graphs," *J. Graph Theory* **25** (1997), 53–58.

[18] W. Gasarch, "A survey of recursive combinatorics," in Handbook of recursive mathematics, Vol. 2, 1041-1176, *Stud. Logic Found. Math.* **139**, North-Holland, Amsterdam, 1998.

[19] M. Goldstern and M. Kojman, "Universal arrow free graphs, " *Acta Math. Hungary* **73** (1996), 319–326.

[20] A. Hajnal and J. Pach, "Monochromatic paths in infinite coloured graphs," in *Finite and infinite sets, Vol. I, II (Eger, 1981),* 359–369, Colloq. Math. Soc. János Bolyai **37,** North-Holland, Amsterdam, 1984.

[21] C. Ward Henson, "Countable homogeneous relational systems and $\aleph_0$-categorical theories," *J. Symb. Logic* **37** (1972), 494–500.

[22] G. Higman, "Ordering by divisibility in abstract algebras," *Proc. London Math. Soc.*, **2**, 326–336, 1952.

[23] W. Hodges, "*Model Theory*," *Encyclopedia of Mathematics and its Applications* **42**, Cambridge University Press, Cambridge, 1993. xiv+772 pp.

[24] S. Huczynska and V. Vatter, "Grid classes and the Fibonacci dichotomy for restricted" permutations, *Electronic J. Combinatorics* **13**, research paper 54, 13 pp., electronic, 2006.

[25] D. Knuth, *The Art of Computer Programming, Vol. 1: Fundamental Algorithms,* Addison-Wesley Publishing Co., Reading, MA, 1969.

[26] P. Komjáth, "Some remarks on universal graphs," *Discrete Math.* **199** (1999), 259–265.

[27] P. Komjáth and J. Pach, "Universal graphs without large bipartite subgraphs," *Mathematika* **31** (1984), 282–290.

[28] P. Komjáth, A. Mekler, and J. Pach, "Some universal graphs," *Israel J. Math* **64** (1988), 158–168.

[29] I. Kříž and R. Thomas, "On well-quasi-ordering finite structures with labels," *Graphs Combin.* **6** (1990), 41–49.

[30] P. Komjáth and J. Pach, "Universal elements and the complexity of certain classes of infinite graphs," *Discrete Math.* **95** (1991), 255–270.

[31] J. B. Kruskal, "Well-quasi-ordering, the tree theorem, and Vazsonyi's conjecture," *Trans. American Math. Soc.* **95** (1950), 210–225.

[32] A. Lachlan and R. Woodrow, "Countable ultrahomogeneous undirected graphs," *Trans. Amer. Math. Soc.* **262** (1980), 51-94.

[33] B. J. Latka, "Finitely constrained classes of homogeneous directed graphs," *J. Symb. Logic* **59** (1994), 124–139.

[34] B. J. Latka, "Structure theorem for tournaments omitting N_5," *J. Graph Theory* **42** (2002), 165–192.

[35] B. J. Latka, "A structure theorem for tournaments omitting $IC_3(I, I, L_3)$," Preprint, 1994. Revised, 2009.

[36] R. Laver, "Well-quasi-orderings and sets of finite sequences," *Math. Proceedings Cambridge Philosophical Society* **79** (1976), 1–10.

[37] S. Mozes, "Tilings, substitution systems and dynamical systems generated by them," *J. d'Analyse Mathématique* **55** (1989), 139–186.

[38] M. Murphy, "*Restricted Permutations, Antichains, Atomic Classes, and Stack Sorting*," Ph. D. thesis, St. Andrews, 2002.

[39] M. Murphy, and V. Vatter, "Profile classes and partial well-order for permutations," (Permutation patterns (Otago, 2003)), *Electron. J. Combin.* **9** (2002/03), Research paper 17, 30 pp. (electronic).

[40] C. St. J. A. Nash-Williams, "On well-quasi-ordering finite trees," *Proc. Cambridge Phil. Soc.* **59** (1863), 833–835.

[41] J. Pach, "A problem of Ulam on planar graphs," *Eur. J. Comb.* **2** (1981), 357–361.

[42] M. Petkovšek, "Letter graphs and well-quasi-order by induced subgraphs," in *Algebraic and topological methods in graph theory (Lake Bled, 1999), Discrete Math.* **244** (2002), 1-3, 375–388.

[43] M. Pouzet, "Un bel ordre d'abritement et ses rapports avec les bornes d'une multirelation," *C. R. Acad. Sci. Paris Sér. A-B* **274** (1972), A1677–A1680.

[44] V. Pratt, "Computing permutations with double-ended queues. Parallel stacks and parallel queues," *Fifth Annual ACM Symposium on Theory of Computing (Austin, Tex., 1973)*, 268–277. ACM, New York, 1973.

[45] R. Rado, "Universal graphs and universal functions," *Acta Arith.* **9** (1964), 331–340.

[46] N. Ruškuc, "Decidability questions for pattern avoidance classes of permutations," in *Third International Conference on Permutation Patterns, Gainesville, Fla., 2005.*

[47] G. Sacks, *Higher Recursion Theory*, Perspectives in Mathematical Logic. Springer-Verlag, Berlin, 1990, xvi+344 pp.

[48] J. Schmerl and W. Trotter, "Critically indecomposable partially ordered sets, graphs, tournaments and other binary relational structures," *Discrete Math* **113** (1993), 191–205.

[49] D. Spielman and M. Bona, "An infinite antichain of permutations," *Electronic J. Combinatorics* **7** (2000), 4 pp. (electronic). Link to "Comments."

[50] R. Tarjan, "Sorting using networks of queues and stacks," *J. Assoc. Comput. Mach.* **19** (1972), 341–346.

[51] S. Torrezão de Sousa and J. K. Truss, "Countable homogeneous coloured partial orders," *Dissertationes Math.* **455** (2008), 48 pp.

[52] S. Waton, *On Permutation Classes Defined by Token Passing Networks, Gridding Matrices and Pictures: Three Flavours of Involvement,* Ph.D. thesis, St. Andrews, 2007.