COMBINATORIAL PROBLEMS CONNECTED WITH FINITE HOMOGENEITY

GREGORY L. CHERLIN

Abstract. Lachlan has developed an elaborate theory of stable homogeneous structures for a finite relational language. Without the stability hypothesis, there is no general theory as yet, and one arrives quickly at quite concrete combinatorial or graph theoretic problems. We state some of these problems and explain their origins in model theory.

Introduction

My main objective here is to explain the following two problems in model theory, and to relate them to more concrete combinatorial problems which can be studied without regard to the model theoretic context. In their most abstract formulations, the problems are as follows.

Problem 1. *Are there uncountably many pseudofinite $\aleph_0$-categorical structures?*

Problem 2. *Is there an effective procedure to decide whether a proposed classification problem arising in the study of homogeneous structures has at most a countable number of solutions?*

Now as stated, these questions involve some notions which can be interpreted in more than one way, and it is necessary to formulate them unambiguously. This will be done in the body of the paper. For the sake of comparison, I will now state two related and far more concrete problems.

Problem 3. *Are there arbitrarily large finite graphs Γ which are triangle-free, and have in addition the 3-extension property: for any two disjoint sets A, B of vertices in Γ with $|A \cup B| = 3$ and A independent (i.e., edgeless), there is a vertex v in Γ joined to all vertices in A and none in B? (Here A or B may be empty.)*

Supported by grant DMS 8603157. This paper was prepared at M. S. R. I., Fall 1989, during the year in logic. My thanks to Hrushovski, Klin, Latka, Mekler, Thomas, and Wood for interesting discussions.

This may not seem like a very tough constraint, but at present the largest known example is the Higman-Sims graph on 100 vertices (pointed out by Simon Thomas).

Problem 4. *For which finite tournaments T is it the case that the set of all finite tournaments which do* not *embed T is well quasi-ordered with respect to the relation of embeddability?*

Some results on this problem have been obtained by my student Brenda Latka.

Problems 3, 4 are not literally special cases of Problems 5, 2, but the natural approaches to those problems do pass through problems like Problem 3 and 4. This will be explained in more detail below.

Contents

1. From homogeneous graphs to homogeneous structures

All graphs or other structures under consideration will be assumed *countable*—possibly finite. A graph is said to be *homogeneous* if any isomorphism between two of its finite induced subgraphs is the restriction of an automorphism. This is the strongest transitivity condition that can be placed on the automorphism group of a graph, and these graphs have been completely classified. The finite homogeneous graphs are classified in [18, 24] and the infinite ones in [32]. There are only three types, up to complementation (if Γ is homogeneous, so is the complementary graph on the same set of vertices):

I. Two exceptional finite primitive graphs: the pentagon, and the edge graph of the complete bipartite graph $K_{3,3}$;
II. $m \cdot K_n$: m disjoint copies of a complete graph, where the parameters m, n may be finite or infinite;
III. Γ_n (Henson): generic omitting K_n (for $n < \infty$), or fully generic (for $n = \infty$).

Here the intuitive meaning of "generic" is: as rich as possible, subject to the given constraint. There are several constructions of Γ_∞, which is known variously as the random graph, or Rado's graph. It is the graph one gets with probability one by selecting edges and nonedges independently with constant edge probability p in $(0, 1)$. It is also the Paley graph over an infinite, locally finite, not quadratically closed field of odd characteristic containing $\sqrt{-1}$. The graph Γ_3 will be described more explicitly in §2.

It is noteworthy that the infinite homogeneous graphs fall into two quite distinct types. Those occurring in Type II are just the infinite members of a natural family of (mostly finite) homogeneous graphs parametrized by extremely simple numerical invariants. Those occurring in the third category do not have finite homogeneous analogs, but the random graph Γ_∞ does have a weaker finite approximation property: if a sequence of random finite graphs of increasing size is selected, then with probability 1 the first order theory of these graphs converges to the theory of Γ_∞ [11, 16]. This classification prompted Lachlan to develop a structure theory for a broad class of homogeneous structures, which we will now sketch.

The notion of homogeneity obviously makes sense in any number of other contexts, including edge-colored graphs, directed graphs, combinatorial geometries, hypergraphs, and also in more algebraic settings like groups, rings, and modules, in which case the finiteness condition occurring in the definition of homnogeneity should be replaced by finite generation.

We will exclude problems arising in the more algebraic settings (which have been intensively studied in concrete cases) by restricting ourselves to *relational structures* for a finite language; that is, no functions are allowed. One would expect that whatever positive results occur in this context would have analogs for *algebraic* systems which are *uniformly locally finite*, and it is worth remarking that this includes all the classical geometries associated with the classical groups over finite fields, but this more general context involves additional problems that we do not wish to consider at the present time. By restricting ourselves to relational systems we remain in the vicinity of graph theory and combinatorics.

Structures which are homogeneous in a category of structures carrying a fixed finite number of relations are called *finitely homogeneous*, and from this point on we use the term "homogeneous" to refer to this notion only. With this terminology, model theorists, with their usual penchant for generality, ask whether there is a structure theory applicable to all homogeneous structures.

The most incisive contributions to such a theory to date are due to Alistair Lachlan. He observed that the graphs occurring in Types I-II are exactly those which are *stable* in the sense of classification theory, a subject which has been primarily concerned with the classification of *uncountable* structures. By developing a finitary analog of this theory, Lachlan [29] obtained a structure theory for the stable homogeneous structures for any fixed finite relational language L (that is, for any fixed category of relational systems) showing that the structures will lie in a *finite* number of families, parametrized by simple numerical invariants; the infinite structures in this classification are just those for which at least one of the associated numerical parameters is infinite.

The notion of stability is quite technical. In the present context it is equivalent to the following. Let us say that a homogeneous structure Γ for a given language L is of *finite type* if every first order sentence true in Γ is true also in a finite homogeneous L-structure. For homogeneous structures, the stability condition is equivalent to being of finite type. (This is one of the results of the structure theory.)

Lachlan's theory therefore includes a classification of all the *finite* homogeneous structures associated with a finite relational language, which is qualitatively quite similar to the one obtained explicitly in the particular case of graphs. His theory also shows that there are only countably many stable homogeneous structures. Of course, there are only countably many homogeneous graphs all told, stable or not, but this remark is rather misleading. There are usually uncountably many homogeneous structures of a given type; for example there are uncountably many homogeneous directed graphs [21].

To put Lachlan's structure theory in the clean form I have just outlined, it is necessary at some point [10] to invoke results of permutation group theory which depend on the classification of the finite simple groups, and make use of detailed information about those groups. One may say that group theory is needed to classify the primitive sections of a large, finite, homogeneous structure, while model theory controls how the pieces may fit together. On the other hand, if the language happens to be a *binary* language—as many of the most natural examples are—the necessary information can be obtained without invoking group theory [31].

Since the original work of Lachlan, improvements have been made on both the group theoretic and the model theoretic side. An idea of Knight and Harrington simplifies the analysis of the effect of a variation of numerical parameters within a given family [25], and the permutation group theory has been much improved by Kantor, Liebeck, and Macpherson [23], using Aschbacher's approach [2].

I intend to focus on problems lying entirely outside the scope of this highly developed and elegant theory. If we drop the stability hypothesis, the theorems have a decidedly anecdotal flavor, and it is very much an open question whether anything more can be said at the level of generality favored by the model theoretic point of view.

The problems mentioned in the introduction have been touched on already in passing during our discussion. Problem 5 concerns the extent to which Lachlan's theory can be extended to structures satisfying a weaker condition of finite approximability than that afforded by stability together with homogeneity. Problem 2 concerns the determination of the point at which countable classifications become impossible. In the next sections we will formulate these questions more precisely.

2. Pseudofiniteness conditions

We are concerned with countable structures, in finite languages, though not always with homogeneous ones. One of the standard finiteness conditions occurring in model theory is the condition of $\aleph_0$-*categoricity*: the structure Γ is $\aleph_0$-categorical if its automorphism group has only finitely many orbits in its natural action on Γ^k, for each k.

There are two conditions of finite approximability we will want to consider. The first is relatively straightforward.

The Finite Model Property

> Every first order sentence true in Γ is true also in some finite structure. In other words, Γ satisfies all sentences which hold in all finite structures.

The second approximability condition is more technical, and will not be examined closely here, but is included for the sake of comparison, as it is relatively well understood at this point.

Smooth Approximation Property

> Every finite subset of Γ is contained in a finite substructure $A \subseteq \Gamma$ such that A smoothly approximates Γ in the sense that $\mathrm{Aut}\, A$ and $(\mathrm{Aut}\, \Gamma)^A$ have exactly the same orbits on A^k for each k. (Here $(\mathrm{Aut}\, \Gamma)^A$ is the group induced on A by its setwise stabilizer in $\mathrm{Aut}\, \Gamma$.)

The following relationships are pertinent:

(1) Stable finitely homogeneous structures are $\aleph_0$-categorical and are smoothly approximable by finite submodels.
(2) $\aleph_0$-categorical structures which are smoothly approximable by finite submodels have the finite model property.

There is a great deal of difference between the finite model property and the condition of smooth approximability. For example, the random graph has the finite model property (cf. §1), but it is not smoothly approximable. In fact the condition of smooth approximability is so restrictive that an explicit countable list of the primitive $\aleph_0$-categorical smoothly approximable structures has been given [**?**] (The term "primitive" is being used here as in permutation group theory, to mean that there are no nontrivial invariant equivalent relations.)

Using the classification in [23] and some further model theoretic argument, one should be able to show by following in the footsteps of [1] and [21] that there are only countably many $\aleph_0$-categorical structures which are smoothly approximable by finite substructures. This has not been carried through as yet.[1] On the other hand one would expect to find $2^{\aleph_0}$ $\aleph_0$-categorical structures with the finite model property.

Problem 5. *Are there uncountably many $\aleph_0$-categorical structures in a finite language with the finite model property?*[2]

[1] It turned out to be complicated.

[2] If one allows a language with relations of unbounded complexity the problem trivializes using random structures.

There are a number of ways one might try to build such structures, the simplest being to construct a suitable family of finitely homogeneous structures. Finitely homogeneous structures are automatically $\aleph_0$-categorical, and there are $2^{\aleph_0}$ of them, as noted in §1; furthermore it may well be that most of the known structures are pseudofinite, in which case Problem 5 should already be settled—but we do not know whether that is the case.

Problem 6. *Are there* $2^{\aleph_0}$ *finitely homogeneous structures with the finite model property?*

(One might even ask if there are uncountably many such structures that are known *not* to have the finite model property, but one can easily find such examples by incorporating a dense linear order into the language.)

A positive solution to this problem also gives a positive solution to Problem 5. There may well be very other approaches to Problem 5.

At present we know of no general method to determine whether or not a specific finitely homogeneous structure has the finite model property, and in fact this question has been settled for only countably many of the known finitely homogeneous structures. We have noted that the stable ones do have the finite model property, as well as the random graph. On the other hand a dense linear order is finitely homogeneous, and without the finite model property.

We can make the problem a great deal more concrete, as follows.

Problem 7. *Are there $2^{\aleph_0}$ homogeneous directed graphs with the finite model property?*

Extending work of Lachlan on homogeneous tournaments [28], I have given a classification of all the homogeneous directed graphs [6, 7, 8]. With countably many exceptions, these directed graphs are all analogs of the Type III homogeneous graphs, namely: generic directed graphs constrained by the condition that they not contain any one of a specified family of tournaments. The analog of a tournament in the undirected case is simply a complete graph, so these are indeed the analogs of the Type III homogeneous graphs.

Thus the meaning of Problem 7 is: are there $2^{\aleph_0}$ generic directed graphs (constrained by the omission of certain tournaments) with the finite model property? As far as I am aware, it is possible that *all* such directed graphs have the finite model property, which would settle the issue. It is also possible that none are (assuming there is at least one forbidden tournament).

As there are only countably many homogeneous graphs, one cannot look for the solution to Problem 5 there. On the other hand, the analog of Problem 7 for graphs remains open, and it seems more natural to first consider that problem for graphs. We formulate the problem explicitly:

Problem 8. *Which graphs of the form Γ_n with n finite have the finite model property?*

The set of all such n is an initial segment of the natural numbers, quite possibly consisting of all natural numbers. The reason for this is quite simple: if $v \in \Gamma_{n+1}$ then the induced graph on the set of neighbors of v is isomorphic to Γ_n, hence approximations to Γ_{n+1} contain approximations to Γ_n. (Nonlogicians may want to insert a few more lines of argument here.)

The first nontrivial case is $n = 3$. We will now state the problem for this case as concretely as possible. We consider the generic triangle-free graph Γ_3. It is characterized up to isomorphism (among triangle-free countable graphs) by the following *extension property*:

> **Ext:** Given two disjoint finite sets of vertices $A, B \subseteq V(\Gamma)$ with A independent (i.e., there are no edges on A), there is a vertex v of Γ connected to all vertices of A, and none of B.

If we restrict our attention to sets A, B with $|A \cup B| \leq k$, we will speak of the *k-extension* property Ext_k. Very easy model theoretic considerations show that Problem 8, for $n = 3$, is entirely equivalent to the following:

Problem 9. *For arbitrary k, are there arbitrarily large finite triangle-free graphs with the k-extension property?*

As mentioned in the introduction, this problem is nontrivial for $k = 3$; the largest currently known triangle-free graph with the 3-extension property is the Higman-Sims graph, on 100 vertices. Norbert Sauer has pointed out that if one takes the n-subsets of a set of size $3n-4$, linking two if their intersection has at most one point, one obtains a limited version of the 3-extension property, subject to the further restriction $|A| \leq 2$.

It is interesting to examine the probabilistic aspects of this question. One could of course try to approximate Γ_n by a large "random" K_n-free graph, that is one would like to prove that for every first order property Φ of Γ_n, the probability that Φ holds in a large finite K_n-free graph goes to 1. Unfortunately there is no known probability model for which this is true. In the simplest model, in which a K_n-free graph is selected at random from the set of all K_n-free graphs on a fixed (large) finite set of vertices, the first order theory of the graph selected does converge (with probability 1) to the theory of a natural graph, but not to the theory of Γ_n; instead the finite K_n-free graphs tend to approach the random n-partite graph [26], which is quite surprising. These graphs are not homogeneous, though if they are viewed as graphs with a partition (given by an equivalence relation) then they become homogeneous in the richer category.

In other words, if we consider the class of finite K_n-free graphs, there are two natural constructions of a finitely homogeneous structure associated with this class:

(1) The generic K_n-free graph, which is homogeneous as a graph but may or may not have the finite model property;
(2) The random n-partite graph, which certainly has the finite model property but is not homogeneous *as a graph*; by chance or otherwise, it *is* homogeneous for a very simple language.

It would be quite interesting to know if this general situation recurs in the context of directed graphs. Fix a finite set of tournaments, and let $\mathcal{A}$ be the set of finite directed graphs in which none of the specified tournaments embed. We know there is an $\mathcal{A}$-generic directed graph $\Gamma(\mathcal{A})$, which is the richest directed graph not embedding any of the specified tournaments; it is easily characterized by extension properties. At the same time there is a probabilistic model: given a finite vertex set V, consider as the sample space all the graphs on the vertex set V which lie in the class $\mathcal{A}$. We would like to know if the first order theories of random graphs from $\mathcal{A}$ over large finite vertex sets converge to a limit theory, the theory of some particular directed graph $\Gamma^*(\mathcal{A})$. If this happens, then $\Gamma^*(\mathcal{A})$ is certainly pseudofinite, and it would then remain to be seen if it is homogeneous for some natural category of relational systems. Granted all of this—a substantial assumption—we would then want to take the closure of the set of complete theories obtained, thereby getting a larger class of pseudofinite theories, and see if all or most of them happen to be $\aleph_0$-categorical (and whether $2^{\aleph_0}$ distinct theories can be obtained in this manner). It is too much too expect that all of this actually works out, but it would be interesting to know what does occur.

It is worth recording the first part of this explicitly.

Problem 10. *Let $\mathcal{T}$ be a finite set of tournaments, and for each N let $S(N)$ be the set of directed graphs Γ on the vertices $\{1, \ldots, N\}$ such that no tournament in $\mathcal{T}$ embeds in Γ. As N tends to infinity, does the theory of a random element of $S(N)$ approach a limit theory, and if so, is that limit theory $\aleph_0$-categorical?*

It is possible that one arrives only at generic n-partite directed graphs in this way.

There is a competing model due to Cameron for selecting a random triangle-free graph [4]. Let S be a complete sum-free symmetric subset of $\mathbb{Z}/N\mathbb{Z}$, that is: $\mathbb{Z}/N\mathbb{Z} = S \sqcup (S+S)$, $S = -S$. Let $\Gamma(S)$ be the graph with vertex set $\mathbb{Z}/N\mathbb{Z}$ and edge relation "$x - y \in S$". Cameron gives a notion of randomly generating a sum-free set, and in contrast to the result of [26] finds that the probability that the result is bipartite is strictly between 0 and 1. Accordingly one might look to this construction as a possible source of triangle-free graphs with the 3-extension property.

3. Classifying classification problems for homogeneous structures

In a few cases, all of the homogeneous structures of a certain type—graphs, directed graphs, and some variants—have been completely classified. One can make quite precise the notion of a classification problem in this context. Such a problem is determined by the following three kinds of data:

— A finite relational language L;
— A *finite* set $\mathcal{A}$ of finite L-structures, called the set of *positive* constraints;
— A *finite* set $\mathcal{B}$ of finite L-structures, called the set of *negative* constraints.

The *solution* of the given problem will be the set of isomorphism types of countable homogeneous L-structures Γ satisfying:

$$\text{Every } A \text{ in } \mathcal{A} \text{ embeds into } \Gamma\text{; no } B \text{ in } \mathcal{B} \text{ embeds into } \Gamma.$$

We emphasize that the *data* defining a classification problem are finite, while the solution set may well be infinite, and even uncountable.

It is natural to distinguish four possible *outcomes* of a classification problem:

- None. There are no solutions.
- Few. There is a positive, finite number of solutions.
- Several. There is a countably infinite set of solutions.
- Many. There are $2^{\aleph_0}$ solutions (the maximum possible number).

As remarked, as a result of work on specific problems, the outcomes of a few classification problems are known.

The general problem, on which the cases treated so far shed little light, would be to say something intelligible about the function

$$\text{Class: Data} \to \text{Outcomes}$$

which takes the data for a given classification problem to one of the four possible outcomes. Since both the inputs and the outputs are finite objects, it then makes sense to ask whether this function is effectively computable. In other words, can one determine the outcome of a classification problem effectively?

The problem may be reformulated as follows. Let (None), (Few), (Several), (Many) denote the sets of data corresponding to the four possible outcomes. The question is whether these sets are *recursive*, that is whether in each of the four cases one can recognize the elements of the given set effectively.

So we state:

Problem 11. *Can one decide effectively whether a classification problem has any solutions at all?*

Problem 12. *Can one decide whether a classification problem has infinitely many solutions?*

Problem 13. *Can one decide whether a classification problem has uncountably many solutions?*

Problem 13 is the precise formulation of the original Problem 2. Problem 11 is a very interesting problem, which has been discussed by Lachlan on various occasions, notably in [30]. He pointed out that Problem 11 provides a way of giving a precise meaning to the problem of classification of arbitrary finitely homogeneous structures, or more exactly the problem of determining whether such a classification is possible. The following terminology seems useful in this connection.

In the first place, there is a natural notion of a *subproblem* of a given problem. For example, the classification of homogeneous partial orders may be viewed as a subproblem of the classification of homogeneous directed graphs. In terms of the data defining the problem, a subproblem is a problem corresponding to an *extension* of the original set of data. (One passes from directed graphs to partial orders by adding negative constraints; one passes from directed graphs to primitive directed graphs by adding positive constraints.) Call a classification problem P *tame* if the set of subproblems of P in (None) is recursive. If one knows how to compute which subproblems of P have solutions, one has a weak classification theorem for the solutions of the problem P (about the weakest imaginable). Problem 11 is to determine whether all classification problems are tame in this sense. If one accepts the identification of the tame problems with the solvable ones, then there is a further question along the same lines:

Problem 14. *Is the set of tame problems recursive?*

This question is of course premature, as it may well be the case that all problems are tame.

Problem 12 appears to have a similar character. If one has a classification which can be viewed as involving finitely many reasonably uniform families of examples, one should be able to handle Problem 12 for all subproblems. One can also imagine handling this problem in less direct ways.

It is instructive to consider Problems 11–13 in the context of homogeneous directed graphs, where as noted earlier we have $2^{\aleph_0}$ examples, and at the same time we have a complete classification. In what follows, I consider these problems under this restriction. In this case Problems 11 and 12 have positive and entirely straightforward solutions.

The homogeneous directed graphs are of two kinds: generic for a class of tournaments, and exceptional. I have no precise definition of the term "exceptional", but in the case at hand we have a very explicit—and of course countable—list of the finitely many families of exceptional homogeneous directed graphs. Thus one can determine easily whether a classification problem for homogeneous graphs with finitely many constraints has any solutions, or infinitely many, in the class of exceptional homogeneous directed graphs. If the answer to one of these questions is negative, it remains only to solve the same problem for the generic examples constrained by tournaments. This goes as follows.

Let the data have positive constraints $\mathcal{A}$, and negative constraints $\mathcal{B}$. If there is a negative constraint $B \in \mathcal{B}$ such that every subtournament of B embeds in one of the positive constraints, our problem has no solutions in the class of directed graphs generic for a set of tournament constraints. Otherwise the problem does have a solution of this type. Next suppose that some negative constraint B has the property that every nonlinear subtournament of B embeds in one of the positive constraints. Then the problem has only finitely many solutions of the specified type. In all other cases there are infinitely many solutions.

When we turn to Problem 13 in the context of homogeneous directed graphs, we encounter a much more difficult problem. The exceptional directed graphs are irrelevant here, and the question reduces to the following.

Problem 15. *Given finitely many constraints on a homogeneous directed graph, can we determine effectively whether they can be met by* $2^{\aleph_0}$ *solutions which are generic for a set of tournament constraints?*

This can be reformulated more concretely as follows. A *well quasi-ordering* is a partial ordering in which every infinite subset contains an increasing sequence. Equivalently, one may impose a pair of conditions: there is no strictly decreasing sequence, and no infinite set of incomparable elements (called an "antichain"). If we consider the partial ordering $\mathcal{P}$ of all finite tournaments, ordered by the relation of embeddability, then $\mathcal{P}$ certainly satisfies the first of these two conditions, but not the second: [21] contains an explicit example of an infinite antichain in $\mathcal{P}$.

Given a finite set of tournaments $\mathcal{T}$, let $\mathcal{P}_{\mathcal{T}}$ be the set of finite tournaments which do not embed any tournament in $\mathcal{T}$. We may reformulate Problem 15 as follows.

Problem 16. *Given a finite set of tournaments $\mathcal{T}$, can one determine effectively whether the set $\mathcal{P}_{\mathcal{T}}$ is well quasi-ordered by the embeddability relation, equivalently: whether $\mathcal{P}_{\mathcal{T}}$ contains an infinite antichain?*

Let us trace through the reduction of Problem 15 to Problem 16. Let $\mathcal{A}$, $\mathcal{B}$ be the positive and negative constraints in a classification problem for homogeneous directed graphs. Given a positive solution to Problem **??**, we can determine effectively whether there is a set $\mathcal{T}$ of finite tournaments with $\mathcal{P}_{\mathcal{T}}$ not well quasi-ordered such that:

- i Every tournament in $\mathcal{T}$ embeds in some negative constraint;
- ii For every negative constraint $B \in \mathcal{B}$, some tournament in $\mathcal{T}$ embeds in B.

If there is, let $\mathcal{X}$ be an infinite antichain in $\mathcal{P}_{\mathcal{T}}$. Let $\mathcal{X}^*$ be the set of tournaments embedding in some positive constraint. We may assume that for every tournament T in $\mathcal{X}^*$, T embeds in all the tournaments in $\mathcal{X}$, or in none of them (shrink $\mathcal{X}$ if need be), and that $\mathcal{X}^* \cap \mathcal{X} = \emptyset$. For any subset $\mathcal{Y}$ of $\mathcal{X}$, let $\Gamma_{\mathcal{Y}}$ be the directed graph which is generic subject to omitting the tournaments of $\mathcal{Y}$. These directed graphs then constitute $2^{\aleph_0}$ distinct solutions to the original problem.

On the other hand, suppose that for every such set $\mathcal{T}$, the partial order $\mathcal{P}_{\mathcal{T}}$ is well quasi-ordered. If Γ is a directed graph which is generic subject to omitting a certain family $\mathcal{X}$ of tournaments, and Γ meets our constraints $\mathcal{A}, \mathcal{B}$, then let $\mathcal{T}$ be the set of tournaments which embed in at least one negative constraint, and do not embed in Γ. The set $\mathcal{T}$ has the properties i, ii above, so $\mathcal{P}_{\mathcal{T}}$ is well quasi-ordered. Let $\mathcal{X}^*$ be the set of minimal elements of $\mathcal{X}$ which do not lie in $\mathcal{T}$; this is an antichain in $\mathcal{P}_{\mathcal{T}}$, hence is finite. Notice that Γ is the generic directed graph omitting $\mathcal{T} \cup \mathcal{X}^*$, hence is determined by a finite set of data. It follows that in this case there are only countably many such directed graphs.

Thus Problem 15 reduces to Problem 16. This is still a difficult problem to attack directly. I will summarize what is known in the case of a single constraint: $|\mathcal{T}| = 1$. Let $\mathcal{Q}$ be the class of finite tournaments T such that $\mathcal{P}_{\{T\}}$ is well quasi-ordered. $\mathcal{Q}$ is downward-closed, is itself well quasi-ordered, and contains all linear tournaments. We wish to know whether $\mathcal{Q}$ is effectively recognizable (recursive). One large class of tournaments *not* in $\mathcal{Q}$ may be described as follows. Say an edge of a tournament is *special* if it embeds in two distinct oriented 3-cycles (nonlinear triangles). Henson's construction [21] shows that any tournament in which the set of special edges contains a cycle (disregarding orientation) is not in $\mathcal{Q}$. At the other extreme, a tournament with *no* special edges embeds in a linearly ordered sequence of oriented 3-cycles, and such a tournament is in $\mathcal{Q}$ if and only if it is either linear, or has at most 4 vertices. This is proved by constructing enough antichains to cover the other cases explicitly. It should be noted that as one proceeds with this sort of analysis, the set of tournaments which "might be" in $\mathcal{Q}$ gradually shrinks toward the well quasi-ordered set $\mathcal{Q}$. If at some point one can show that the set of remaining candidates is reduced to a well quasi-ordered set, it would then follow that our problem (for $|\mathcal{T}| = 1$) *is* decidable, as the "true" $\mathcal{Q}$ would then necessarily be determined by finitely much additional data. Thus there is a real possibility that one might prove that a problem of this type is effectively decidable while having no actual decision procedure.

Model theorists will recognize the preoccupation with the "countable vs. $2^{\aleph_0}$" frontier as a formal analog of Shelah's classification of first order theories by the behavior of their spectrum function at larger cardinals. Shelah's point was that there are a finite number of criteria which determine on which side of the gap a particular theory falls. We are asking a similar question here: can the corresponding gap, for countable homogeneous structures, be recognized effectively. In a second phase, we would like a particular kind of structure theory applicable to classification problems having only countably many solutions, perhaps in terms of a natural coding by elements of standard well quasi-ordered sets, like the set of words in a finite alphabet under embedding, or some more elaborate set of finite labelled trees. On the other side, classification problems with $2^{\aleph_0}$ solutions which happen to have explicit solutions are reminiscent of the "deep" theories in Shelah's classification. There cannot be any very real connection, but it is not terribly surprising that formal analogies appear in contexts which have a certain family resemblance.

It should be added that our formulation of Problems 11–13 is not quite right, unless the answer is positive in all cases. Otherwise, the question should be rephrased: determine the precise level of complexity (in the recursion theoretic sense) of the following sets of classification problems:

- (None) As before, problems with no solutions;
- (Finite) Problems with finitely many solutions;
- (Countable) Problems with countably many solutions.

In this formulation, (None)$\subseteq$(Finite)$\subseteq$(Countable).

As Lachlan pointed out, (None) is certainly r.e., by Fraissé's theory of amalgamation classes. In the same vein, the collection of problems with at most one solution is Π^0_2, (Finite) is Σ^0_3, and (Countable) is Π^1_1. These estimates are made by writing out the definitions; they are the best ones I know. In Problems 11–13, we ask whether all these sets are recursive.

4. Outrageous conjectures, algebraic closure, the small index property, and reducts

To round out this account I would like to mention some other issues arising in connection with finitely homogeneous structures.

4.1. **Outrageous conjectures.** Because we know so little about finitely homogeneous structures in general, it is possible to formulate some fairly outrageous conjectures consistent with the evidence, which would trivialize aspects of Problem 2. We might as well begin with the most outrageous one of all.

Problem 17. *Is it true that all homogeneous binary structures are known?*

Though in principle fairly easy to demolish, this conjecture is by definition always consistent with the evidence. We want to experiment with variants which offer a more promising target.

Problem 18. *Classify the finite homogeneous binary structures.*

I have been informed that Gol'fand has strong partial results on this problem which have not yet been published. The special case of finite primitive homogeneous binary structures is of considerable interest; they should be quite rare, but as far as I know a definitive classification has not been proved. Of course Lachlan's theory applies here, saying that for each finite binary language a list is available in principle, as the result of a certain effective computation. We now concern ourselves exclusively with the infinite case, and for simplicity we confine ourselves to primitive structures.

Problem 19. *Is every primitive infinite binary symmetric homogeneous structure generic for a free amalgamation class?*[3]

In the case of graphs, the Type III graphs are those which are generic with respect to a certain notion of free amalgamation. The precise definition involves Fraissé's notion of an amalgamation class, and will not be given here, but the question is whether the extreme phenomenon encountered in the case of graphs (in which every example is either

[3]Homogeneous metric spaces refute this. But one may modify the question so as to allow for any specified family of forbidden triangles, replacing "free amalgamation" by a canonical amalgamation procedure for the resulting class. This is actually plausible.

finite, imprimitive, or Type III) is valid for all homogeneous structures in symmetric binary languages, that is: for all edge-colored graphs with a finite number of colors. In asymmetric structures, the principal known departure from this pattern is encountered in the context of partial orderings, and variations on this theme.

We have certain possibilities for generating new homogeneous structures from old. Let L_1 be a binary language, construed as a set of 2-types with an action of $\mathbb{Z}/2\mathbb{Z}$ (which takes a type to its converse, fixing the symmetric types). Let L be a language with a surjection $f : L \twoheadrightarrow L_1$, where maps between languages are taken to commute with the action of $\mathbb{Z}/2\mathbb{Z}$. Let Γ_1 be a structure on which the algebraic closure operator is trivial (compare the §4.2). Then there is a unique homogeneous L-structure Γ, the generification of Γ_1 along f, such that its finite substructures are exactly the finite L-structures whose collapse along f is a substructure of Γ_1.

In a related vein, let L_1, L_2 be two languages with a common sublanguage L_0, and let Γ_1, Γ_2 be homogeneous L_1-, L_2-structures respectively, so that the reducts of Γ_1, Γ_2 to L_0 are isomorphic to a fixed homogeneous structure Γ_0. Assume that there is an amalgamation procedure for substructures of Γ_0 which is compatible with amalgamation procedures for substructures of both Γ_1 and Γ_2. Let L be the language:

$$\Delta(L_0) \sqcup (L_1 \backslash L_0) \times (L_2 \backslash L_0) \subseteq L_1 \times L_2,$$

with ΔL_0 the diagonal in $L_1 \times L_2$, and let $f_i : L \to L_i$ be given by projection. Then there is a unique homogeneous L-structure Γ, the free join of Γ_1, Γ_2 over L_0, whose finite substructures are exactly the finite L-structures whose collapses along f_1, f_2 are substructures of Γ_1, Γ_2 respectively.

The simplest case of the last construction has $L_0 = \{=\}$, and the algebraic closure operator trivial in both Γ_1 and Γ_2. In this case we just call Γ the free join of Γ_1 and Γ_2.

Problem 20. *Is every primitive infinite binary homogeneous structure either freely generated or else one of a countable list of classifiable exceptions?*

This question lacks precision, but this is of little importance: the answer is in any case negative. Taking the free join of a linear order and a freely generated homogeneous structure provides $2^{\aleph_0}$ examples of homogeneous binary structures which are not freely generated, as Hrushovski pointed out to me. (That this sort of thing is useful has been noticed before in more interesting situations by Schmerl.) It does not seem too useful to try to reformulate the problem to take this into account; we retreat to Problem 19, and eventually we hope to be driven all the way back to Problem 11, the most interesting version.

4.2. **Algebraic closure and pseudoplanes.** Let Γ be finitely homogeneous, $A \subseteq \Gamma$ finite, and let $\operatorname{Aut}\Gamma_A$ be the pointwise stabilizer of A in $\operatorname{Aut}\Gamma$. The *definable* closure $\operatorname{dcl}(A)$ is the fixed point set of $\operatorname{Aut}\Gamma_A$; the *algebraic* closure $\operatorname{acl}(A)$ is the union of the finite orbits of $\operatorname{Aut}\Gamma_A$.

Problem 21. *If Γ is infinite, primitive, binary, and finitely homogeneous, is $acl(A) = A$ for all finite A?*

The following remark is due to Carol Wood. Let Γ be the graph whose vertices are the two element subsets of a fixed infinite set, with two vertices joined by an edge if disjoint. Let Γ^* be Γ expanded to a ternary structure by adjoining the relation $R(x, y, z)$ defined by: "The sets x, y, z have an element in common". Then Γ^* is primitive and homogeneous for a ternary language, and the algebraic closure of a set of two unlinked vertices is nontrivial. So the restriction to the binary case is necessary.

We can generalize Problem 21 to the imprimitive case as follows.

Problem 22. *Let Γ be homogeneous for a finite binary language, $a \in$* $\operatorname{acl}(A)$*. Does it follow that there are equivalence relations E_i which are* $\operatorname{Aut} M$*-invariant, and elements $a_i \in A$, such that $E_i(a, a_i)$ holds for all i, and the set thus defined:*

$$\{x : \textit{For all } i,\ E_i(x, a_i) \textit{ holds}\}$$

is finite?

Taken as a conjecture, Problem 22 is stronger, and seems somewhat riskier, than Problem 21. The main point here is that by generalizing the problem, we get a statement which may conceivably have an inductive proof.

There is a variation on Problem 21 which is not restricted to the binary case (and where we also allow imprimitivity).

Problem 23. *Is there a finitely homogeneous pseudoplane in Lachlan's sense* [27]*?*

Simon Thomas has shown (unpublished) that no pseudoplane is homogeneous for a finite binary language. After Hrushovski's construction of an $\aleph_0$-categorical pseudoplane, the notion of a finitely homogeneous pseudoplane seems less outrageous; but it still seems outrageous.

We emphasize that we have no efficient way of acquiring a large body of examples of homogeneous binary structures, hence no strong empirical basis for anticipating a positive answer to Problem 21.

4.3. **The small index property.** The abstract groups $\mathrm{Aut}\,\Gamma$ are relatively inscrutable objects from an elementary group theoretic point of view. Mati Rubin [34] has a beautiful argument showing that in many cases the structure Γ can be recovered from the abstract group Γ. There is a natural topology on $\mathrm{Aut}\,\Gamma$, the topology of pointwise convergence for the discrete topology on Γ. The main point is to recover the topology on $\mathrm{Aut}\,\Gamma$ from its group structure. In certain cases this can be done quite simply. We say that Γ has the *small index* property if every subgroup of $\mathrm{Aut}\,\Gamma$ of countable index is open. This then implies that the topology on Γ is determined by the group structure. One expects stable finitely homogeneous structures to have this property (cf. [15]).

Problem 24. *Do all finitely homogeneous structures have the small index property?*

This is open even in the case of the random graph Γ_∞ of §1! Rubin's work does show that in this case the group structure on $\mathrm{Aut}\,\Gamma_\infty$ *does* determine the topology in some fashion, but casts no light on the more specific Problem 24.

4.4. **Reducts.** A reduct of a structure Γ is a closed subgroup of $\mathrm{Sym}(\Gamma)$ containing $\mathrm{Aut}\,\Gamma$; it is *nontrivial* if it is not the full symmetric group (so $\mathrm{Aut}\,\Gamma$ counts as nontrivial). The study of reducts was suggested by Cameron, I believe, and actually launched by Simon Thomas in Spring 1989 with the following theorem.

Fact 4.1 ([38])**.** Γ_∞ *has exactly 4 nontrivial reducts.*

The following description of these reducts is simple, but not very enlightening. Let P_n be the n-place relation on Γ_∞ defined as follows:

There are an odd number of edges on the points $x_1, \ldots, x_n$.

Let $\Gamma_\infty \restriction P_n$ be the structure whose elements are the vertices of Γ_∞ and whose only relation is the relation P_n. The reducts of Γ_∞ are the groups $\mathrm{Aut}(\Gamma_\infty \restriction P_n)$ for $n > 1$; this sequence is periodic modulo 4.

Thomas' proof is an elegant application of the Nešetril-Rödl extension [33] of Ramsey's theorem to the context of the random graph. The same argument shows that the Γ_n have no nontrivial proper reducts for n finite. Thomas proposes:

Problem 25. *Does a finitely homogeneous structure always have finitely many reducts?*

The most manageable portion of this problem would be the case of a finitely homogeneous structure corresponding to a freely generated amalgamation class.

Appendix A. Problems (and variants)

Problem (Problem 5). *Are there uncountably many pseudofinite $\aleph_0$-categorical structures?*

Problem (Problem 6). *Are there $2^{\aleph_0}$ finitely homogeneous structures with the finite model property?*

Problem (Problem 7). *Are there $2^{\aleph_0}$ homogeneous directed graphs with the finite model property?*

Problem (Problem 8). *Which graphs of the form Γ_n with n finite have the finite model property?*

Problem (Problem 9). *For arbitrary k, are there arbitrarily large finite triangle-free graphs with the k-extension property?*

Problem (Problem 10). *Let $\mathcal{T}$ be a finite set of tournaments, and for each N let $S(N)$ be the set of directed graphs Γ on the vertices $\{1, \ldots, N\}$ such that no tournament in $\mathcal{T}$ embeds in Γ. As N tends to infinity, does the theory of a random element of $S(N)$ approach a limit theory, and if so, is that limit theory $\aleph_0$-categorical?*

Problem (Problem 2). *Is there an effective procedure to decide whether a proposed classification problem arising in the study of homogeneous structures has at most a countable number of solutions?*

Problem (Problem 11). *Can one decide effectively whether a classification problem has any solutions at all?*

Problem (Problem 12). *Can one decide whether a classification problem has infinitely many solutions?*

Problem (Problem 13). *Can one decide whether a classification problem has uncountably many solutions?*

Problem (Problem 14). *Is the set of tame problems recursive?*

Problem (Problem 15). *Given finitely many constraints on a homogeneous directed graph, can we determine effectively whether they can be met by $2^{\aleph_0}$ solutions which are generic for a set of tournament constraints?*

Problem (Problem 16). *Given a finite set of tournaments $\mathcal{T}$, can one determine effectively whether the set $\mathcal{P}_{\mathcal{T}}$ is quasi well-ordered by the embeddability relation, equivalently: whether $\mathcal{P}_{\mathcal{T}}$ contains an infinite antichain?*

Problem (Problem 17)**.** *Is it true that "all" homogeneous binary structures are known?*

Problem (Problem 18)**.** *Classify the finite homogeneous binary structures.*

Problem (Problem 19)**.** *Is every primitive infinite binary symmetric homogeneous structure generic for a free amalgamation class?*

Problem (Problem 20 [refuted, see §4])**.** *Is every primitive infinite binary homogeneous structure either freely generated or else one of a countable list of classifiable exceptions?*

Problem (Problem 21)**.** *If* Γ *is infinite, primitive, and homogeneous for a finite binary relational language, is* $acl(A) = A$ *for all finite* A*?*

Problem (Problem 22)**.** *Let* Γ *be homogeneous for a finite binary relational language,* $a \in \mathrm{acl}(A)$*. Does it follow that there are equivalence relations* E_i *which are* $\mathrm{Aut}\, M$*-invariant, and elements* $a_i \in A$*, such that* $E_i(a, a_i)$ *holds for all* i*, and the set thus defined:*

$$\{x : \textit{For all } i,\ E_i(x, a_i) \textit{ holds}\}$$

is finite?

Problem (Problem 23)**.** *Is there a finitely homogeneous pseudoplane in Lachlan's sense* [27]*?*

Problem (Problem 24)**.** *Do all finitely homogeneous structures have the small index property?*

Problem (Problem 25)**.** *Does a finitely homogeneous structure always have finitely many reducts?*

References

[1] G. Ahlbrandt and M. Ziegler, "Quasi finitely axiomatizable totally categorical theories," *Ann. Pure Appl. Logic* **30** (1986), 63-82.
[2] M. Aschbacher, "On the maximal subgroups of the finite classical groups," *Invent. Math.* **76** (1984), 469-514.
[3] C. Ash, "Undecidable $\aleph_0$-categorical theories," *Notices Amer. Math. Soc.* **18** (1971), 423.
[4] P. Cameron, "On the structure of a random sum-free set," *Prob. Th. Rel. Fields* **76** (1987), 523-531.
[5] G. Cherlin, "Homogeneous tournaments revisited," *Geom. Ded.* **26** (1988), 231-240.
[6] 1 in3 pt, "Homogeneous n-tournaments," submitted.
[7] 1 in3 pt, "Homogeneous directed graphs I. The imprimitive case," *in* **Logic Colloquium 1985**, ed. Paris Logic Group, North-Holland, 1987, 67-88.

[8] 1 in3 pt, "Homogeneous directed graphs II. Graphs omitting I_∞, submitted.
[9] 1 in3 pt, "Homogeneous directed graphs II. Graphs embedding I_∞, submitted.
[10] G. Cherlin and A. Lachlan, "Stable finitely homogeneous structures," *Trans. Amer. Math. Soc.* **296** (1986), 815-851.
[11] Yu. Glebski, D. Kogan, M. Liogon'kiĭ, and V. Talanov, "Volume and fraction of satisfiability of formulas of the lower predicate calculus," *Kibernetika* **2** (1969), 17-27.
[12] Ya. Gol'fand and M. Klin, "On k-homogeneous graphs," *in* **Algorithmic studies in combinatorics**, 76-85, Nauka, Moscow, 1978.
[13] J. Dixon, P. Neumann, and S. Thomas, "Subgroups of small index in infinite symmetric groups, *Bull. London Math. Soc.* **18** (1986), 580-586.
[14] A. Ehrenfeucht, "There are continuum $\aleph_0$-categorical theories," *Bull. Acad. Polon. Sci.* **20** (1972), 425-427.
[15] D. Evans, W. Hodges, and I. Hodkinson, "Automorphisms of bounded abelian groups," preprint.
[16] R. Fagin, "Probabilities on finite models," *J. Symb. Logic* **41** (1976), 50-58.
[17] R. Fraissé, "Sur l'extension aux relations de quelques propriétés des ordres," *Ann. Ecole Norm. Sup.* **7** (1954), 361-388.
[18] A. Gardiner, "Homogeneous graphs and stability," *J. Austr. Math. Soc. Ser. A* **21** (1976), 371-375.
[19] A. Gardiner, "Homogeneous graphs," *J. Comb. Th.* **20** (1976), 94-102.
[20] W. Glassmire, Jr., "There are $2^{\aleph_0}$ countably categorical theories," *Bull. Acad. Polon. Sci.* **19** (1971), 185-190.
[21] C. W. Henson, "Countable homogeneous relational systems and categorical theories," *J. Symb. Logic* **37** (1972), 494-500.
[22] E. Hrushovski, "Totally categorical structures," *Trans. Amer. Math. Soc.* **313** (1989), 131-159.
[23] W. Kantor, M. Liebeck, D. Macpherson, "$\aleph_0$-categorical structures smoothly approximated by finite substructures," preprint.
[24] M. Klin, *Trudy Nikolaev Korablestoit. Inst.* **87** (1974), 3-8.
[25] J. Knight, and A. Lachlan, "Shrinking, stretching, and codes for homogeneous structures," preprint.
[26] Ph. Kolaitis, H. Prömel, and B. Rothschild, "K_{l+1}-free graphs: asymptotic structure and a 0-1 law," *Trans. Amer. Math. Soc.* **303** (1987), 637-671.
[27] A. Lachlan, "Two conjectures on the stability of ω-categorical theories," *Fund. Math.* **81** (1974), 133-145.
[28] 1 in3 pt, "Countable homogeneous tournaments," *Trans. Amer. Math. Soc.* **284** (1984), 431-461.
[29] 1 in3 pt, "On countable stable structures which are homogeneous for a finite relational language," *Israel J. Math* **49** (1984), 69-153.
[30] 1 in3 pt, "Homogeneous structures," *in* **Proceedings of the ICM 1986**, American Mathematical Society, 1987, 314-321.
[31] A. Lachlan and S. Shelah, "Stable structures homogeneous for a finite binary language," *Israel J. Math.* **49** (1984), 155-180.
[32] A. Lachlan and R. Woodrow, "Countable ultrahomogeneous undirected graphs," *Trans. Amer. Math. Soc.* **262** (1980), 51-94.
[33] J. Nešetril and V. Rödl, "The Ramsey property for graphs with forbidden complete subgraphs," *J. Comb. Th.* **20** (1976), 243-249.

[34] M. Rubin, "On the reconstruction of $\aleph_0$ categorical structures from their automorphism groups," preprint, 1987.
[35] J. Schmerl, "Countable homogeneous partially ordered sets," *Alg. Univ.* **9** (1979), 317-321.
[36] J. Sheehan, "Smoothly embeddable subgraphs," *J. London Math. Soc.* **9** (1974), 212-218.
[37] S. Shelah, **Classification Theory and the Number of Non-isomorphic Models**, North-Holland, Amsterdam, 1978.
[38] S. Thomas, "Reducts of the random graph," preprint.
[39] J. Truss, "The group of the countable universal graph," *Math. Proc. Cambridge Phil. Soc.* **98** (1985), 213-245.

Rutgers University