Discrete Mathematics 95 (1991) 49–67
North-Holland

The age of a relational structure

Peter J. Cameron

School of Mathematical Sciences, Queen Mary and Westfield College, Mile End Road, London, UK, E1 4NS

Received 3 October 1989
Revised 26 January 1990

Abstract

Cameron, P.J., The age of a relational structure, Discrete Mathematics 95 (1991) 49–67.

The age of a relational structure is the class of its finite substructures. This paper surveys aspects of the relationship between structures and their ages, concentrating on three areas: universality, $\aleph_0$-categoricity and homogeneity; measure and Baire category applied to the class of countable structures with given age; and enumerative aspects (the profile of an age).

1. Introduction

In algebra and combinatorics (though not usually in topology), an object is determined by information about its finite subsets. Examples include whether or not the product of two given group elements is equal to a third, whether two vertices of a graph are adjacent, whether an element of an ordered set is less than another.

The age of a structure is the class of its finite substructures (up to isomorphism). This paper is a survey of results about ages of infinite relational structures. I will always assume that the infinite structures under discussion are countable. (In interesting cases, nothing is lost by this assumption, as we shall see).

The topics addressed are:

(i) How do we recognise an age? Of structures with a given age, is there one which is in some way distinguished, e.g. by properties of symmetry, universality or axiomatisability?

(ii) Can we 'measure' in some way the structures whose age is contained in a given age, so as to give meaning to the statement 'most such structures have property $\mathscr{P}$'?

(iii) How does the function enumerating isomorphism types of n-element structures in an age behave, and which functions can be realised? How do properties of the structure X affect this?

As these questions suggest, there are many and varied links with logic, combinatorics, permutation groups, measure theory, and topology.

2. Ages and universality

A *relational structure* consists of a set X with a family $R = \{r_\alpha : \alpha \in A\}$ of finitary relations on X. That is, r_α is a subset of X^n, where $n = n(\alpha) \geq 1$ is the arity of r_α.

With any relational structure (X, R) is associated a language which has relation 'symbols' corresponding to the relations in R. As a notational convenience, I will not distinguish between relations and relation symbols. The language also contains the standard paraphernalia of first-order logic (variables, connectives, quantifiers, and parentheses), so that well-formed formulae can be built up in the usual way.

A *substructure* of (X, R) is a relational structure (Y, S), where Y is a subset of X and $S = \{s_\alpha : \alpha \in A\}$, $s_\alpha = r_\alpha \cap Y^n$, $n = n(\alpha)$. The relations correspond to those in R, and strictly speaking different symbols should be used, as here; but the notational convention above suggests using the same notation for corresponding relations. Of course, this is common mathematical practice; when we write $x < y$, or $x + y = z$, we do not specify whether the arguments of the relations lie in the natural numbers, integers, rationals or reals.

Most objects in combinatorics (graphs, directed or undirected, total or partial orders, equivalence relations) and algebra (groups, rings, etc.) can be regarded as relational structures. A group, for example, is described by a ternary relation which holds at (x, y, z) precisely when $xy = z$. Note that substructures in this case are much more general than subgroups; any subset carries a substructure! (If this situation is undesirable, the usual procedure is to allow function and constant symbols in the language. Then, however, most of what is said here will not apply.)

One small technical point must be mentioned. Often, especially in mathematical logic, relations of arity 0 are permitted; I have forbidden them here. A relation of arity n can be thought of as Boolean function on X^n; so a relation of arity 0 is just a Boolean variable, which may be true or false in a particular situation. If this were allowed, then the empty set would carry several different relational structures, which would be inconvenient. This problem does not arise in logic, where it is customary to forbid the empty set as the domain of a structure (to avoid paradoxes like 'nothing is better than a good book; a bad book is better than nothing; so . . .').

The domain of a relational structure in this paper will always be finite or countable. Again, following common practice, the name of the domain is often applied to the structure.

The *age* of the relational structure X, written $\mathrm{Age}(X)$, is the class of all finite structures which are isomorphic to substructures of X. The fact that the age, as thus defined, is a proper class rather than a set does not really give any trouble; but it will often be convenient to assume that the domain of a structure in the age consists of the first n natural numbers, for some n. (In combinatorial terms, we are thinking of Age(X) as a class of labelled structures).

The terminology is due to Fraïssé (see [13]).

Knowing the age of X obviously gives us some information about X itself. We can tell, for example, whether X is a graph, whether directed or undirected, whether with or without loops, whether complete or not, bipartite or not, etc.; whether it is a totally or partially ordered set; and so on. (In all these cases, the appropriate axioms are universal statements involving only finitely many points). How much further information we get depends very sensitively on the particular case. If Age(X) consists of the finite complete graphs, then X is uniquely determined as the countably infinite complete graph. On the other hand, if Age(X) consists of all finite totally ordered sets, then X is any countable totally ordered set, and there are $2^{\aleph_0}$ possibilities for it. Note that, in either case, Age(X) contains a unique n-element structure (up to isomorphism).

Let Y be a relational structure with the same language as X. We say that Y is *younger than* X if $\mathrm{Age}(Y) \subseteq \mathrm{Age}(X)$. Clearly, any substructure of X is younger than X. The converse is false. (A two-way infinite path, and the disjoint union of three one-way infinite paths, have the same age; but neither is a substructure of the other.) We say that X is *universal* if every structure younger than X is isomorphic to a substructure of X. (In the terminology of Komjáth and Pach, in their article in this volume, such an X is universal in the class of structures younger than X. In special cases, e.g. universal graphs (Rado [31]), the unqualified word is already standard. Fraïssé's term is *rich for its age.*)

Cantor's celebrated theorem [6] asserts, in part, that Q (as totally ordered set) is universal. We will generalise this in various ways later.

Note that there are usually many universal structures with a given age, since any structure with a universal substructure is itself universal.

We now turn to the question: How do you recognise an age?

A class $\mathcal{K}$ of finite structures has the *joint embedding property* if for any B_1, $B_2 \in \mathcal{K}$, there exists $C \in \mathcal{K}$ and embeddings $g_i : B_i \to C$ for $i = 1, 2$. That is, any two members of $\mathcal{K}$ can be simultaneously embedded in a larger member of $\mathcal{K}$. This condition, together with trivial book-keeping requirements, is necessary and sufficient for $\mathcal{K}$ to be an age.

Theorem 2.1. *The class $\mathcal{K}$ of finite structures is the age of a countable structure X if and only if $\mathcal{K}$ is closed under isomorphism, closed under taking substructures, contains only countably many non-isomorphic members, and has the joint embedding property.*

Proof. The conditions on $\mathcal{K}$ are trivially necessary; so assume that they hold. Enumerate the isomorphism types in $\mathcal{K}$, say $A_0, A_1, \ldots$, and define a sequence $B_0, B_1, \ldots$ recursively as follows: $B_0 = \emptyset$; B_{n+1} is a structure in $\mathcal{K}$ jointly embedding B_n and A_n, $n \geq 0$.

We can regard B_n as a substructure of B_{n+1}. The let $X = \bigcup_{n\geq 0} B_n$. Then every A_n is a substructure of X; and any finite substructure of X is a substructure of B_n for some n, and hence is contained in $\mathscr{K}$ (since $B_n \in \mathscr{K}$ and $\mathscr{K}$ is substructure-closed). □

Remark. All conditions except the third are satisfied by the age of a relational structure of arbitrary cardinality. So, if the countability condition is satisfied, (and, in particular, if the language is finite), then there exists a countable structure with the same age. This justifies our restriction to countable structures.

3. First-order axioms

The theorem of Cantor alluded to above characterises $\boldsymbol{Q}$ as countable dense totally ordered set without endpoints. Thus $\boldsymbol{Q}$ is uniquely determined by the assumption of countability together with a collection of axioms, sentences (of first-order logic) in the language of ordered sets.

A structure X with this property is called $\aleph_0$-categorical. Thus the countable structure X is *$\aleph_0$-categorical* if there is a set Σ of sentences in the language of X such that any countable model of Σ is isomorphic to X. Note that countability has to be required as a 'nonlogical axiom' here: the upward Löwenheim-Skolem Theorem says that, if Σ has an infinite model, then it has arbitrarily large infinite models. See Chang and Keisler [7].

A remarkable characterisation of $\aleph_0$-categorical structures was found by Engeler [9], Ryll-Nardzewski [32] and Svenonius [35]. In this result, the language need not be assumed relational; functions and constants are permitted.

Theorem 3.1. *For the countable structure X, the following are equivalent*:
(i) *X is $\aleph_0$-categorical;*
(ii) *for every n, there are only finitely many n-types over* Th(X);
(iii) *for every n, the automorphism group of X has only finitely many orbits on X^n.*

Here an n-type is a maximal set of formulae in n variables consistent with Th(X), the theory of X (the set of all sentences true in X); that is, everything that can be said in first-order language about an n-tuple of points in a structure with the same theory. More is true: if X is $\aleph_0$-categorical, then every n-type over Th(X) is realised in X (i.e. there is an n-tuple in X satisfying it), and two n-tuples lie in the same orbit of Aut(X) if and only if they satisfy the same n-type.

The surprising fact is that axiomatisability of X is equivalent to the existence of a very large automorphism group. But perhaps, bearing in mind Felix Klein's philosophy (as expressed in the Erlanger Programm), this is not so unexpected.

As a simple consequence, we see that, if X is $\aleph_0$-categorical, then the number of n-element structures in Age(X) is finite for all n. So, if X has infinitely many relations of arity at most n for some n, then most of them are 'redundant' (all are Boolean combinations of a finite number).

Problem. Find necessary and sufficient conditions for a class $\mathscr{K}$ to be the age of an $\aleph_0$-categorical structure. (The class of all finite graphs whose connected components are paths is not the age of any $\aleph_0$-categorical structure).

The theorem has a pretty corollary, which is presumably known to the experts, though I am not aware of a reference.

Theorem 3.2. *An $\aleph_0$-categorical structure is universal.*

Proof. Let X be $\aleph_0$-categorical, and let Y be younger than X; assume that the domain of Y is $\boldsymbol{N}$, the natural numbers. Let F_n be the substructure of Y on $\{0, \ldots, n-1\}$. We call two embeddings $f_1, f_2: F_n \to X$ equivalent if there is an automorphism g of X such that $f_1 g = f_2$. By Theorem 3.1 and the fact that Y is younger than X, there are at least one but only finitely many equivalence classes for each n.

Construct a tree as follows: The nodes at level n are the equivalence classes of embeddings of F_n in X. Nodes on levels n and $n+1$ are adjacent in the tree if they have representatives f, f' respectively such that f is the restriction of f' to F_n.

König's Infinity Lemma [19] guarantees that the tree has an infinite path, starting at the root. It is easy to see that we can choose representatives of the nodes on this path consistently (that is, so that the embedding of F_n is the restriction of that of F_{n+1}). Now the union of all these embeddings is clearly a $1-1$ map from Y into X; and it is an isomorphism from Y onto a substructure of X, since an instance of a relation holds in Y if and only if it holds in F_n for some n. □

This is one generalisation of Cantor's Theorem. We will soon see another.

4. Homogeneity

The ordered set $\boldsymbol{Q}$ has a very strong symmetry property: whenever $x_1 < \cdots < x_n$ and $y_1 < \cdots < y_n$, there is an automorphism of $\boldsymbol{Q}$ carrying $(x_1, \ldots, x_n)$ to $(y_1, \ldots y_n)$. (Interpolate in the intervals (x_i, x_{i+1}) by linear functions, and at the two ends by appropriate shifts). Thus, $\boldsymbol{Q}$ is homogeneous, in the following sense: A structure X is *homogeneous* if, given any isomorphism $f: A \to B$ between finite substructures of X, there is an automorphism g of X whose restriction to A is f.

This is the strongest symmetry condition we can impose on a relational structure. Fraïssé [12] proved an important theorem which is both a characterisation and a construction of homogeneous structures, bearing a striking resemblance to Theorem 2.1.

A class $\mathscr{K}$ of finite structures has the *amalgamation property* if, whenever we have structures $A, B_1, B_2, C \in \mathscr{K}$ and $f_i : A \to B_i$ are embeddings ($i = 1, 2$), there exist $C \in \mathscr{K}$ and embeddings $g_i : B_i \to C$ ($i = 1, 2$) so that the diagram

$$\begin{tikzcd} A \arrow[r, "f_1"] \arrow[d, "f_2"] & B_1 \arrow[d, "g_1"] \\ B_2 \arrow[r, "g_2"] & C \end{tikzcd}$$

commutes. In other words, any two structures with a common substructure can be jointly embedded so that their intersection contains at least this common substructure. By taking $A = \emptyset$, we see that the amalgamation property implies the joint embedding property. (This, however, depends on our conventions about the empty set.)

Theorem 4.1. *The class $\mathscr{K}$ of finite structures is the age of a countable homogeneous structure X if and only if $\mathscr{K}$ is closed under isomorphism, closed under taking substructures, contains only countably many non-isomorphic members, and has the amalgamation property. Moreover, if these conditions are fulfilled, then the countable homogeneous structure X is unique.*

The unique X is called the *Fraïssé limit* of the class $\mathscr{K}$.

For example, the class of all finite graphs satisfies the hypotheses; its Fraïssé limit was constructed explicitly by Rado [31], and implicitly earlier by Erdős and Rényi [11]. For reasons which will become clear, it is called the random graph. Henson [15] observed that, for any $n \geqslant 2$, the class $H(n)$ of graphs containing no complete subgraph of size n satisfies the hypotheses of Theorem 4.1. (If amalgamation is performed by adding no superfluous edges, then no new K_n will be constructed). The Fraïssé limits of these classes give further homogeneous graphs. Lachlan and Woodrow [22] proved that this is essentially all.

Theorem 4.2. *A countable homogeneous graph is one of the following*:

(a) *a disjoint union of m complete graphs of size n, where at least one of m and n is infinite*;

(b) *complement of* (a) (*a complete multipartite graph*);

(c) *the Fraïssé limit of the class of K_n-free graphs, $n \geqslant 3$*;

(d) *complement of* (c);

(e) *the random graph.*

This major theorem has been applied in Ramsey theory, see Nešetřil and Rödl [27]. Other classes of relational structures in which all homogeneous members have been determined include tournaments (Lachlan [20]) and partial orders (Schmerl [33]).

A subsidiary result in the proof of Theorem 4.1 is useful in its own right.

Theorem 4.3. *The countable structure X is homogeneous if and only if it has the following property: if $A, B \in \mathrm{Age}(X)$ and $f: A \to B$, $h: A \to X$ are embeddings, then there is an embedding $g: B \to X$ such that $fg = h$. It suffices to require this when $|B| = |A| + 1$.*

If X is homogeneous and $\mathrm{Age}(X)$ has only finitely many n-element members for all n, then $\mathrm{Aut}(X)$ has only finitely many orbits on X^n for all n, since isomorphic structures lie in the same orbit. By Theorem 3.1, X is $\aleph_0$-categorical. But we can seen this more directly. Let X be a homogeneous structure with age $\mathcal{K}$. Then a countable structure Y is isomorphic to X if and only if it satisfies the following 'axioms':

(I) For each $A \in \mathcal{K}$, the statement that there exists a substructure of Y isomorphic to A. (This is a first-order existential sentence; the collection of such sentences ensures that $K \subseteq \mathrm{Age}(Y)$.)

(II) For each n, the statement that every n points of Y carry a substructure isomorphic to some member of $\mathcal{K}$. (This ensures that $\mathrm{Age}(Y) \subseteq \mathcal{K}$. It is not in general first-order; but if $\mathcal{K}$ has only finitely many n-element substructures, it is a disjunction over all of these, universally quantified).

(III) For each embedding $f: A \to B$ between members of $\mathcal{K}$, the statement of Theorem 4.3. (This is an $(\forall\exists)$ sentence, i.e. of the form $(\forall x_1 \cdots x_n)(\exists y)\phi$, where ϕ says 'if the substructure on $x_1, \ldots, x_n$ is A, then the substructure on $x_1, \ldots, x_n, y$ is B'.)

Note that (I) and (II) alone assert that $\mathrm{Age}(Y) = \mathcal{K}$.

We see that the axioms for a homogeneous $\aleph_0$-categorical structure can be taken to be $(\forall\exists)$ sentences (where one or the other of the types of quantifiers can be omitted). This implies, in particular, that such structures are model-complete (see [7]).

Another consequence of Theorem 4.3 is the following.

Theorem 4.4. *A homogeneous structure is universal.*

Proof. Let X be homogeneous, and Y younger than X, with the domain of Y equal to $\mathbb{N}$. For each n, let F_n be the substructure of Y on $\{0, \ldots, n-1\}$. By Theorem 4.3, any embedding of F_n in X can be extended to an embedding of F_{n+1}. Doing this for all n and taking the union gives an embedding of X. □

For more on homogeneity and universality, see the article by Komjáth and Pach in this collection.

One final remark. If X is homogeneous, then everything about the embedding of a finite structure A in X (in particular, any first-order formula satisfied by the points in its image) is determined by the isomorphism type of A. A theory is said to admit *quantifier-elimination* if any formula is equivalent, modulo the theory, to a formula without quantifiers. This suggests a relation between homogeneity and quantifier-elimination which is made precise in the following result.

Theorem 4.5. *A theory Σ admits quantifier-elimination if and only if any model of Σ has an elementary embedding into a homogeneous model. In particular, a $\aleph_0$-categorical theory admits quantifier-elimination if and only if its unique countable model is homogeneous.*

A slight variant on the amalgamation property has been studied. The *strong amalgamation property* is satisfied by $\mathscr{K}$ if, in the conclusion of the amalgamation property, we can assume that no unexpected identifications are made; that is, if $b_1 \in B_1$, $b_2 \in B_2$ satisfy $b_1g_1 = b_2g_2$, then there exists $a \in A$ such that $af_1 = b_1$ and $af_2 = b_2$.

Theorem 4.6. *Let X be a countable homogeneous structure. Then* Age(X) *has the strong amalgamation property if and only if the subgroup of* Aut(X) *fixing any finite tuple of points has no additional fixed points.*

The condition holds in $\boldsymbol{Q}$, for example: the orbits of the stabiliser of a tuple are the open intervals into which $\boldsymbol{Q}$ is divided by the points of the tuple. For an application of this concept (in mathematical psychology), see Cameron [2]. Proofs of the theorems in this section appear in the forthcoming book [3].

5. Absolute ubiquity

Informally, I will refer to a structure X as 'ubiquitous' if 'most' (in some sense) of the structures younger than X are actually isomorphic to X. It is quite remarkable, both that this often occurs, and that the property correlates well with other properties of interest. In particular, X is *absolutely ubiquitous* if every structure having the same age as X is isomorphic to X. We observed earlier that the complete graph is absolutely ubiquitous.

Conditions (I) and (II) preceding Theorem 4.4 ensure that a structure Y has prescribed age $\mathscr{K}$. Over a finite relational language, these conditions are first-order, as we noted. Now X is absolutely ubiquitous if and only if they (together with the assumption of countability) characterise X up to isomorphism. So, over a finite relational language, an absolutely ubiquitous structure is $\aleph_0$-categorical, and so has a large group of automorphisms. Just how large was made clear by Hodkinson and Macpherson [16], who gave what amounts to a complete description of these structures.

Theorem 5.1. *Let X be absolutely ubiquitous over a finite relational language. Then there is a partition of X into finitely many parts $Y_1, \ldots, Y_m$ such that the direct product of the symmetric groups on $Y_1, \ldots, Y_n$ is a group of automorphisms of X.*

Some of the parts may be finite. The content of the theorem is that whether or not an n-tuple satisfies a relational depends only on the parts of the partition containing its points and which pairs of points are equal. So a finite amount of information specifies the structure. The full automorphism group may be larger; consider, for example, a complete bipartite graph. A fairly typical absolutely ubiquitous graph is the disjoint union of an infinite complete bipartite graph, an infinite star, and a finite graph.

It follows easily from Theorem 5.1 that, if X is a relational structure over a finite language with the property that every structure younger than X is isomorphic to X, then X is trivial (i.e. $\mathrm{Aut}(X)$ is the symmetric group).

We could also consider properties $\mathcal{P}$ of X for which every structure with the same age as X satisfies $\mathcal{P}$. (Such properties might be called 'absolutely forced', but this term has a different meaning in logic.) The isomorphism type of an absolutely ubiquitous structure satisfies this condition. Clearly, any property of X expressed by a universal or existential sentence satisfies the condition; for example, the properties of being a graph, or of having no triangles.

6. The tree of an age

To investigate ages further, we provide them with some structure.

We take the 'normalised' version of $\mathrm{Age}(X)$, where the domain of any member consists of the first n natural numbers. We make $\mathrm{Age}(X)$ into a tree in the obvious way: nodes at level n are the n-element structures; nodes F_n, F_{n+1} at levels n, $n+1$ are adjacent if the restriction of F_{n+1} to $\{0, \ldots, n-1\}$ is F_n.

Now let $\mathrm{Younger}(X)$ be the set of structures than X, similarly 'normalised' so that their domains are all $\boldsymbol{N}$. Any structure $Y \in \mathrm{Younger}(X)$ defines a path in the tree $\mathrm{Age}(X)$, whose node at level n is the substructure of Y on $\{0, \ldots, n-1\}$. Conversely, any path in $\mathrm{Age}(X)$ specifies uniquely a member of $\mathrm{Younger}(X)$. Thus $\mathrm{Younger}(X)$ is identified with the set of infinite paths in the tree $\mathrm{Age}(X)$.

In the next two sections, we explore two ways of explicating the idea of 'most members of $\mathrm{Younger}(X)$', involving Baire category and measure respectively.

7. Baire category

Think of following a path in $\mathrm{Age}(X)$ as successively specifying more and more information about an infinite structure. Two paths are close together if a lot of information is required to distinguish them. Accordingly, define a metric on Younger(X) (regarded as a set of paths rather than a set of structures) by the rules

$$d(Y, Y) = 0;$$

$$d(Y_1, Y_2) = \tfrac{1}{2}^n \quad \text{if the paths } Y_1, Y_2 \text{ agree to level } n \text{ but no further.}$$

(The actual choice of values is unimportant; any decreasing sequence of positive numbers tending to 0 would do.) The metric space axioms are easily verified; indeed, in placé of the triangle inequality, we have the stronger *ultrametric inequality*

$$d(Y_1, Y_3) \leqslant \max\{d(Y_1, Y_2), (Y_2, Y_3)\},$$

which implies, that, if two open balls meet, then one contains the other.

This metric space is complete: a Cauchy sequence of paths has the property that its members agree on longer and longer initial segments; these initial segments define a path, which is the limit of the sequence.

The set of paths containing a fixed node F is an open ball, and every open ball has this form. Thus, a set S of paths is open if and only if, for every path $Y \in S$, there is a node F on Y with the property that every path containing F is in S. Similarly, a set S of paths is dense if and only if every node lies on a path in S.

A subset of a metric space is *residual* (or *comeagre*, or *the complement of a set of the first category*) if it contains a countable intersection of open dense sets. Residual sets are regarded as 'large'. The Baire category theorem asserts that, in a complete metric space, a residual set is non-empty. Of course, this immediately implies the apparently stronger statement that the intersection of countably many residual sets is non-empty.

Proving the Baire category theorem in our special case is much simpler than in the general case. Let $S_0, S_1, \ldots$ be open and dense. Define a sequence of nodes $F_0, F_1, \ldots$ by the rules: $F_0 = \emptyset$; F_{n+1} is a node above F_n with the property that every path through F_{n+1} lies in S_n (this exists because S_n is open and dense). Then the unique path containing all these nodes lies in $\bigcap_{n \geqslant 0} S_n$.

A slight modification of the argument shows that a residual set is dense. Given any node F, replace the start of the above recursion with $F_0 = F$; the resulting path passes through F.

Now we say that a structure X is *ubiquitous in category* if the set of structures isomorphic to X is residual in Younger(X); and a property $\mathcal{P}$ of X is *forced in category* if a residual subset of Younger(X) satisfies $\mathcal{P}$. (This relation between the theory of X and first-order sentences forced in category is known to logicians as finite forcing, following Abraham Robinson; but the present name seems more appropriate here.)

Theorem 7.1. (i) *Any* ($\forall\exists$) *sentence true in X is forced in category by X.*

(ii) *A homogeneous structure is ubiquitous in category.*

Proof. (i) Let the sentence ϕ be true in X, where ϕ is expressed in the form $(\forall x_1 \cdots x_n)(\exists y_1 \cdots y)\theta(x_1, \ldots, x_n, \quad y_1, \ldots, y_m)$. (For short we write $(\forall \bar{x})(\exists \bar{y})\theta(\bar{x}, \bar{y})$.) The set of structures in Younger(X) satisfying it is the intersection, over all choices of $a_1, \ldots a_n \in \boldsymbol{N}$, of the sets $S(\bar{a}) = S(a_1, \ldots, a_n)$ of structures satisfying $(\exists \bar{y})\theta(\bar{a}, \bar{y})$. So it suffices to show that $S(\bar{a})$ is open and dense. The openness is clear: if $Y \in S(\bar{a})$, so that $\theta(\bar{a}, \bar{b})$ holds in Y for some $\bar{b}(b_1, \ldots, b_m)$, and $k = \max\{\bar{a}, \bar{b}\}$, then any structure whose restriction to $\{0, \ldots, k\}$ agrees with that of Y is also in $S(\bar{a})$. To show denseness, suppose that the substructure F on $\{0, \ldots, k\}$ has been defined for some k, and that $a_1, \ldots, a_n \in F$. Find a copy F' of F inside X; let $a'_1, \ldots, a'_n$ be the corresponding points of X. Since ϕ holds in X, there exist $b'_1, \ldots, b'_m \in X$ such that $\theta(\bar{a}', \bar{b}')$ holds. Then extend F to a structure F'' on an initial segment of $\boldsymbol{N}$ isomorphic to $F|_{\{a'_1, \ldots, a'_n, b'_1, \ldots, b'_m\}}$. Any infinite structure extending F'' is in $S(\bar{a})$.

(ii) We observed in Section 4 that conditions (I)–(III) before Theorem 4.4 characterise the homogeneous structure X. Of these, (I) and (III) comprise countably many $(\forall\exists)$ sentences. (II), which may not even be first-order, simply ensures that $Y \in$ Younger(X), and so is not required here. Thus the isomorphism type of X is a countable intersection of residual sets in Younger(X), and so is residual. □

It follows from Theorem 7.1 that, for any structure X, almost all (i.e. a residual set of) structures younger than X actually have the same age as X, since this is ensured by condition (I) alone.

This is not the end of the story, however. Consider the two-way infinite path X. Age(X) consists of all finite unions of finite paths. It is easily checked that the following property is forced in category by X: *Any finite set is contained in the interior of a path.* Moreover, this property characterises X within Younger(X). So X is ubiquitous in category, though it is certainly not either homogeneous or $\aleph_0$-categorical. The above property, though not first-order, has the general flavour of a $(\forall\exists)$ sentence.

This example is instructive. There are two 'distinguished' structures having the same age: a single infinite path, and the disjoint union of countably many infinite paths. The first is ubiquitous in category, the second universal. They also have contrasting model-theoretic properties. (They have the same theory; they are the prime and saturated countable models of this theory respectively.)

8. Measure

In this section, I consider an alternative interpretation of 'most', using the notion of measure, which is closely related and in a sense dual to that of category (see Oxtoby [29]). We assign a positive real number $m(F)$ to each node F of the tree Age(X), this number being interpreted as the measure of the open ball consisting of paths containing F. We must require that $m(F)$ is equal to the sum of the values of $m(F')$ over all nodes F' immediately above F. (For convenience I assume in this section that the number of n-element structures in Age(X) is finite for all n; this ensures that the sum in question is finite). Any function m satisfying these conditions defines in a standard way a measure on the σ-algebra of Borel subsets of Younger(X).

If the measure of the whole space (i.e. $m(\emptyset)$) is 1, as it will be in what follows, then m is a probability measure. The interpretation is that, in a random member of Younger(X), the probability that the substructure induced on $\{0, \ldots, n-1\}$ is F is equal to $m(F)$.

There are many functions satisfying this requirement. For example, we may take $m(\emptyset) = 1$, and then divide the measure of any node F equally among its children. However, a further condition is desirable:

(+) For any $F \in$ Age(X) with $|F| = n$, and any distinct $x_0, \ldots, x_{n-1} \in \mathbf{N}$, the measure of the set $\{Y \in$ Younger(X): the map $i \mapsto x_i$ $(i = 0, \ldots, n-1)$ is an embedding of F in $Y\}$ is equal to $m(F)$.

This condition is equivalent to the requirement that if F is isomorphic to F' then $m(F) = m(F')$. It is not satisfied by the measure defined above by equal division. Consider, for example, the age $H(3)$ consisting of all triangle-free graphs. The graph with vertex set $\{0, 1, 2\}$ and edges $\{0, 2\}$ and $\{1, 2\}$ has measure $\frac{1}{8}$, while that with edges $\{0, 1\}$ and $\{0, 2\}$ has measure $\frac{1}{6}$.

There is, however, a 'natural' way to attempt to define a measure, which (if it succeeds) ensures that (+) holds. This is motivated by the interpretation of probability as limiting frequency. Given a node F in Age(X), we define

$$m(F) = \lim_{n\to\infty} \frac{\text{number of nodes at level } n \text{ above } F}{\text{number of nodes at level } n}.$$

Informally, the probability that the substructure on $\{0, \ldots, n-1\}$ is F is equal to the limiting frequency of F among all n-element substructures of large labelled structures in Age(X). For example, if X is a graph and F an edge, $m(F)$ is the limiting average density of edges in large labelled subgraphs of X.

Problem. Does this limit exist for any choice of the relational structure X (over a finite language) and any finite substructure F of X?

(Of course, it is easy to construct trees in which the limit does not exist. But, for the comparatively few ages in which I have been able to decide the question, I have found no counterexamples.)

If $m(F)$, as defined above, exists for all $F \in$ Age(X), it is straightforward to show that it is nonnegative, is additive over children of any node, and satisfies (+); so it defines a measure on Younger(X) satisfying all our requirements. Now, much as before, we say that a property $\mathcal{P}$ of X is *forced in measure* if it holds on a set of measure 1 (the complement of a null set), and that X is *ubiquitous in measure* if its isomorphism type is forced in measure.

In the case where Age(X) consists of all finite graphs, the measure is precisely that defined by the informal statement 'choose edges independently with probability $\frac{1}{2}$'. Now the result of Erdős and Rényi [11] alluded to earlier is that (in the terminology introduced earlier) the random graph is ubiquitous in measure—hence its name. The ordered set Q is also ubiquitous in measure. This can be derived from a general result.

Theorem 8.1. *Let X be a countable homogeneous structure whose age has the following property: there is a function f on the natural numbers such that any n-element structure* Age(X) *is contained in exactly $f(n)$ $(n+1)$-element structures. Suppose further that $\sum 1/f(n)$ diverges. Then X is ubiquitous in measure.*

For Q, the hypotheses of this theorem hold, with $f(n) = n + 1$. Ubiquity in measure has some pleasant consequences.

Theorem 8.2. *If X is ubiquitous in measure, then the substructure of X induced on any cofinite subset is isomorphic to X.*

This is proved by showing that the induced measure on the complement of a finite set corresponds to the original measure for structures on that set. The result shows that a structure ubiquitous in measure is sufficiently tough that removal of finitely many points leaves it unchanged. It is very likely true that ubiquity in measure implies other similar toughness results, perhaps even the existence of automorphisms; but little work has been done on this.

There are, however, some surprises. One remarkable fact is that the age $H(3)$ consisting of the finite triangle-free graphs forces in measure the property of being bipartite. (This is the content of a theorem of Erdös, Kleitman and Rothschild [10], according to which the proportion of triangle-free graphs on n vertices which are nonbipartite tends to 0 as $n \to \infty$. From this it follows that $m(F) = 0$ for any nonbipartite F, i.e. the probability that the subgraph on a given finite set is nonbipartite is equal to zero. Since there are only countably many finite subsets, and since a countable union of null sets is null, the random triangle-free graph almost surely has no nonbipartite finite subgraph).

It can be shown that the measure associated with $H(3)$ (which, from the above, is the same as that associated with the age consisting of all finite bipartite graphs) has the following simple description: Given a finite set F, divide it into two subsets by assigning points to one or the other independently with probability $\frac{1}{2}$; then choose edges independently with probability $\frac{1}{2}$ from the pairs with one point in each subset.

There is a countable bipartite graph B which is forced in measure by either of the ages described. (In particular, B is ubiquitous in measure). B is *almost* *homogeneous*: any isomorphism between finite subgraphs which respects the bipartition in B is induced by an automorphism of B. B is also ubiquitous in category.

By contrast, Henson's homogeneous triangle-free graph is ubiquitous in category but not in measure.

9. Profiles

The profile of an age is the function on $\mathbf{N}$ which enumerates the isomorphism types of n-element structures in the age for each n. Throughout this section, we suppose that all values of the profile are finite (this holds, in particular, if the language is finite). The principal question is: which functions from $\mathbf{N}$ to $\mathbf{N}$ occur as profiles? Two specialisations of this question restrict attention to homogeneous structures, or structures over finite languages (for example, graphs).

If X is homogeneous, then the number of isomorphism types of n-element substructures of X is equal to the number of orbits of $\mathrm{Aut}(X)$ on n-element subsets of X.

The fundamental result is as follows.

Theorem 9.1. *The profile of an age is a nondecreasing function.*

Two entirely different proofs are known using linear algebra and Ramsey theory respectively. Both give additional information. For the first proof, let X be an infinite set, and k a positive integer. Let V_k denote the vector space of functions from the set of k-element subsets of X to $\mathbf{Q}$. For $k \leqslant l$, define a linear map θ from V_k to V_l as follows: for $f \in V_k$, $L \subseteq X$, $|L| = l$, set

$$(f\theta)(L) = \sum \{f(K) \colon K \subseteq L, |K| = k\}.$$

Theorem 9.2. *With the above notation,* $\mathrm{Ker}\,\theta = \{0\}$.

Now suppose that the numbers of isomorphism types of k- and l-element structures are n_k, n_l respectively. Let A be the $n_k \times n_l$ matrix whose (i, j) entry is the number of k-sets of the ith isomorphism type contained in an l-set of the jth isomorphism type. Then A represents the restriction of θ to the subspace spanned by the characteristic functions of the isomorphism types of k-sets; so A has rank n_k, and $n_l \geqslant n_k$. (In fact, this argument is valid also for finite sets X, provided that $|X| \geqslant k + l$.)

The second proof uses an extension of Ramsey's Theorem (which, however, is proved using Ramsey's Theorem).

Theorem 9.3. *Suppose that the k-subsets of the infinite set X are coloured with n colours, all of which are used. Then there exist an ordering $(c_1, \ldots, c_n)$ of the colours, and infinite subsets $X_1, \ldots, X_n$ of X, such that for each value of i, X_i contains a set of colour c_i but none of colour c_j for $j > i$.*

Now colour the k-subsets of X with $n = n_k$ colours according to their isomorphism types. Then, for $l \geq k$, at least n different l-sets can be distinguished by the isomorphism types of their k-subsets. This actually shows that the matrix A of the previous proof has its principal $n_k \times n_k$ submatrix upper triangular, if rows and columns are suitably ordered. (This proof also works for sufficiently large finite sets; but here, the lower bound obtained from the proof of a finite version of Theorem 9.3 is an iterated Ramsey number.)

How fast does the profile grow? The next result is an amalgam of results of Pouzet [30] and Macpherson [23].

Theorem 9.4. *Let f be the profile of a binary relational structure. Then either:*
(a) *f grows like a polynomial (that is, there exist $k \in \mathbf{N}$ and $c, d \in \mathbf{R}$ with $c > 0$, such that $cn^k \leq f(n) \leq dn^k$ for all n); or*
(b) *$f(n) \geq \exp(n^{1/2-\epsilon})$ for all sufficiently large n.*

All cases here occur. For a disjoint union of complete graphs of size $k+1$, the profile grows like a polynomial of degree k. For a disjoint union of infinitely many infinite complete graphs, the profile is the partition function, which is very roughly $\exp(n^{1/2})$ (see Hall [14]).

Pouzet showed that either (a) holds or f grows faster than any polynomial. His proof involves a delicate analysis of the sub-ages of an age, as partially ordered set. Macpherson showed that, in the latter case, it is possible to encode partitions of n into substructures of X of size cn for some fixed c, using Ramsey's Theorem.

Problem. Suppose that f is a profile which is *subexponential*, that is, $f(n) < \exp(n^{1-c})$ for sufficiently large n, where $c > 0$. Is it true that there exists $p \in \mathbf{N}$ such that, for any $\epsilon > 0$ and all sufficiently large n,

$$\exp(n^{p/(p+1)-\epsilon}) < f(n) < \exp(n^{p/(p+1)+\epsilon})?$$

(There are some reasons for thinking that such a strange spectrum of growth rates might be true, and examples having such growth are known for any $p \in \mathbf{N}$).

Considerably more detailed results are known for homogeneous structures. If X is homogeneous, then isomorphism types of n-subsets are orbits of $\mathrm{Aut}(X)$; our blanket assumption in this section just says that X is $\aleph_0$-categorical. On the other hand, given any permutation group G, there is a homogeneous relational structure X such that $G \leq \mathrm{Aut}(X)$ and G has the same orbits on X^n as $\mathrm{Aut}(X)$ (for all n). So results in the present case are equivalent to results about orbits of permutation groups, and are often phrased in this way.

We call X *transitive* if no nontrivial unary relation is definable (without parameters) in X, and *primitive* if no nontrivial equivalence relation is definable. (These conditions hold precisely when the permutation group $G = \mathrm{Aut}(X)$ is transitive, resp. primitive, in the usual sense). Macpherson [25] showed the following.

Theorem 9.5. *There is a constant $c > 1$ such that, if X is a primitive homogeneous structure with profile f, then either $f(n) = 1$ for all n, or $f(n) > c^n$ for all sufficiently large n.*

The proof is a considerable elaboration of his argument for Theorem 9.4, encoding trees rather than partitions. Examples where the profile grows no faster than exponentially are quite rare; most of them are 'treelike' objects (see [1]). It would be interesting to know more about the spectrum of possible 'exponential constants', i.e. values of $\lim_{n\to\infty} f(n)^{1/n}$, for profiles f of primitive homogeneous structures (and, in particular, whether this limit always exists).

Macpherson [26] has also found connections between growth rates faster than exponential for primitive homogeneous structures and model-theoretic properties such as stability.

The growth rate of the profile can be pushed up artificially by having relations in the language of arbitrarily large arity. However, for finite relational languages, there is an upper bound.

Theorem 9.6. *If X is a structure over a finite relational language L, then the profile of X is bounded by the exponential of a polynomial depending only on L.*

This follows by observing that, if $m_1, \ldots, m_k$ are the arities of the relations in L, then the total number of n-element L-structures is

$$2^{n^{m_1} + \cdots + n^{m_k}}.$$

Note that, by Fraïssé's Theorem 4.1, this bound is attained by a homogeneous structure over L.

There are structures with growth rates intermediate between these levels. Two examples follow.

(i) Consider the age consisting of finite sets carrying two independent total orders. (There is a homogeneous structure with this age.) The profile is the factorial function.

(ii) The class of finite line graphs is the age of the line graph of the infinite complete graph. Its profile grows a little faster than exponentially, but the precise asymptotic behaviour, surprisingly, seems to be unknown (despite detailed analysis by Wright [38], among others).

10. Further directions

One of the most important recent developments, on which I have touched only incidentally, is the connection with a class of model-theoretic concepts which had been introduced in connection with the question: given a theory, how many models of given infinite cardinality can it have? These concepts, such as stability, the strict order property, and the independence property, can be found in Shelah [34]. Stable $\aleph_0$-categorical infinite structures can be regarded as 'limits' of finite ones, and their behaviour is closely controlled by their finite substructures. One of the most striking results is that of Lachlan [21], which gives a kind of classification of all finite or stable infinite homogeneous structures over any finite relational language. The ω-stable, $\aleph_0$-categorical structures are quite well understood, due to work of Zil'ber [39], Cherlin, Harrington and Lachlan [8], and Hrushovski [17].

For structures of large infinite cardinality, the profile can be extended to infinite cardinal numbers. The second proof of Theorem 9.1 shows immediately that the number of countable structures is at least as great as the number of n-element substructures (if this is finite). Results on higher cardinalities have been obtained by Kierstead and Niykos, Pouzet and Woodrow [37], and (for orbits of permutation groups) by Neumann [28]. But the best result, due to Macpherson, Mekler and Shelah [26], extends the monotonicity of the profile to infinite cardinal numbers sufficiently much smaller than $|X|$. It is thought that the profile is nondecreasing for all cardinals below $|X|$. This is known to fail at $|X|$, however. Consider $X = \omega_1$, the first uncountable ordinal. The substructures of X are all the ordinals up to ω_1. There are thus uncountably many countable substructures (up to isomorphism), but only one of size $\aleph_1$.

The concepts of Baire category and measure, which were little more than descriptive tools in this paper, have been used in their more traditional roles in existence proofs. (To show that some object has property $\mathcal{P}$, show that almost all do.) These proofs mainly involve showing that certain structures admit groups of automorphisms with certain properties. Here is a single example. Among countable graphs which admit a cyclic automorphism, almost all (a residual set of measure 1) are isomorphic to the random graph R. Hence R admits cyclic automorphisms, and indeed admits $2^{\aleph_0}$ nonconjugate cyclic automorphisms. See Cameron and Johnson [4] for further results in this direction.

The generating function of a profile turns out to be a specialisation of a Pólya-type *cycle index function* associated with the age. Another specialisation of the same series is the exponential generating function for labelled structures in the age. In the case of a homogeneous structure X, this cycle index function can be associated with its automorphism group; the two specialisations mentioned count orbits of this group on n-sets and on n-tuples of distinct points respectively. There are rules for computing the cycle indices of direct and wreath products and of point stabilisers, and links with many topics in enumerative combinatorics (e.g.

Cameron and Taylor [5]). The cycle index itself can be regarded as a kind of specialisation of a 'combinatorial formal power series' in the sense of Joyal [18] (who would regard an age as a *species*).

The philosophy here is to approximate infinite structures by finite ones. This can be extended to approximate automorphisms of infinite structures by partial automorphisms of finite ones. For example, Truss (see [36] for a survey) has shown that in many cases (for example, the random graph), there is a 'generic' or residual conjugacy class of automorphisms, which have infinitely many cycles of each finite length (but, surprisingly, no infinite cycles). Further connections with permutation groups will appear in the forthcoming book [3].

Finally, I refer to Fraïssé's book [13] on the theory of relations, which contains a treasure trove of information in many areas related to the topic of this survey.

References

[1] P.J. Cameron, Some treelike objects, Quart. J. Math. Oxford Ser. (2) 38 (1987) 155–183.
[2] P.J. Cameron, Groups of order-automorphisms of the rational numbers with prescribed scale type, J. Math. Psych. 33 (1989) 163–171.
[3] P.J. Cameron, Oligomorphic Permutation Groups, London Math. Soc. Lecture Notes (Cambridge Univ. Press, Cambridge, 1991).
[4] P.J. Cameron and K.W. Johnson, An investigation of countable Burnside groups, Math. Proc. Cambridge Philos. Soc. 102 (1987) 223–231.
[5] P.J. Cameron and D.E. Taylor, Stirling numbers and affine equivalence, Ars Combin. 20 (B) (1985) 3–14.
[6] G. Cantor, Beiträge zur Begründung der transfiniten Mengenlehre, Math. Ann. 46 (1895) 481–512.
[7] C.C. Chang and H.J. Keisler, Model Theory (North-Holland, Amsterdam, 1973).
[8] G.L. Cherlin, L.A. Harrington and A.H. Lachlan, $\aleph_0$-categorical, $\aleph_0$-stable structures, Ann. Pure Appl. Logic 28 (1985) 103–135.
[9] E. Engeler, Äquivalenzklassen von n-Tupeln, Z. Math. Logik Grundlag. Math. 5 (1959) 126–131.
[10] P. Erdős, D.J. Kleitman and B.L. Rothschild, Asymptotic enumeration of K_n-free graphs, Colloq. Internat. Teorie Combin. II (1977) 19–27.
[11] P. Erdős and A. Rényi, Asymmetric graphs, Acta Math. Acad. Sci. Hungar. 14 (1963) 295–315.
[12] R. Fraïssé, Sur certains relations qui généralisent l'ordre des nombres rationnels, C. R. Acad. Sci. Paris 237 (1953) 540–542.
[13] R. Fraïssé, Theory of Relations (North-Holland, Amsterdam, 1986).
[14] M. Hall Jr, Combinatorial Theory (Blaisdell, Waltham, MA 1967).
[15] C.W. Henson, A family of countable homogeneous graphs, Pacific J. Math. 38 (1971) 69–83.
[16] I.M. Hodkinson and H.D. Macpherson, Relational structures determined by their finite induced substructures, J. Symbolic Logic 53 (1988) 222–230.
[17] E. Hrushovski, to appear.
[18] A. Joyal, Une théorie combinatoire des series formelles, Adv. in Math. 42 (1981) 1–82.
[19] D. König, Theorie der Endlichen und Unendlichen Graphen (Chelsea, New York, 1950).
[20] A.H. Lachlan, Countable homogeneous tournaments, Trans. Amer. Math. Soc. 284 (1984) 431–461.
[21] A.H. Lachlan, On countable stable structures which are homogeneous for a finite relational language, Israel J. Math. 49 (1984) 69–153.
[22] A.H. Lachlan and R.E. Woodrow, Countable ultrahomogeneous undirected graphs, Trans. Amer. Math. Soc. 262 (1980) 51–94.

[23] H.D. Macpherson, Growth rates in infinite graphs and permutation groups, Proc. London Math. Soc (3) 51 (1985) 285–294.
[24] H.D. Macpherson, Orbits of infinite permutation groups, Proc. London Math. Soc. (3) 51 (1985) 246–284.
[25] H.D. Macpherson, Infinite permutation groups of rapid growth, J. London Math. Soc. (2) 35 (1987) 276–286.
[26] H.D. Macpherson, personal communication, 1989.
[27] J. Nešetřil, personal communication, 1988.
[28] P.M. Neumann, Homogeneity of infinite permutation groups, Bull. London Math. Soc. 20 (1988) 305–312.
[29] J.C. Oxtoby, Measure and Category (Springer, Berlin, 1971).
[30] M. Pouzet, Application d'une propriété combinatoire des partes d'un ensemble aux groupes et aux relations, Math. Z. 150 (1976) 117–134.
[31] R. Rado, Universal graphs and universal functions, Acta Arith. 9 (1964) 393–407.
[32] C. Ryll-Nardzewski, On the categoricity in power $\aleph_0$, Bull. Acad. Polon. Sci. Ser. Math. 7 (1959) 545–548.
[33] J. Schmerl, Countable homogeneous partially ordered sets, Algebra Universalis 9 (1979) 317–321.
[34] S. Shelah, Classification Theory and the Number of Non-isomorphic Models (North-Holland, Amsterdam, 1978).
[35] L. Svenonius, $\aleph_0$-categoricity in first-order predicate calculus, Theoria 25 (1959) 82–94.
[36] J.K. Truss, Infinite simple permutation groups—a survey, to appear.
[37] R.E. Woodrow, personal communication, 1986.
[38] E.M. Wright, The number of connected sparsely-edged graphs, III: Asymptotic results, J. Graph Theory 4 (1980) 393–407.
[39] B.I. Zil'ber, The structure of models of uncountably categorical theories, Internat. Congr. Math. 1 (1984) 359–368.