

The random graph and its friends

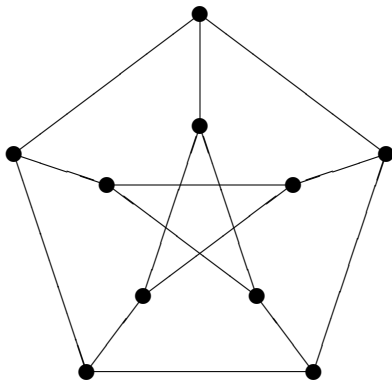
Peter J. Cameron
University of St Andrews

Mathematics Colloquium
University of Vienna
22 March 2017



Example: the Petersen graph

This is the most famous finite graph; a whole book has been devoted to it.



I am going to tell you about the most famous *infinite* graph ...

The random graph

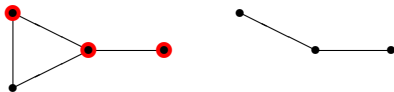
The countable random graph is one of the most extraordinary objects in mathematics.

As well as graph theory and probability, we can turn to set theory (the Skolem paradox) or number theory (quadratic reciprocity, Dirichlet's theorem) for constructions of this object, logic (\aleph_0 -categoricity), group theory (simple groups, Cayley graphs), Ramsey theory (Ramsey classes of structures) or topological dynamics (extreme amenability) for some of its properties, and topology (the Urysohn space) for a related structure.

I will tell you some of its story.

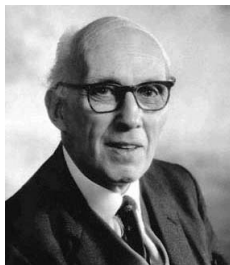
Graphs and induced subgraphs

A **graph** consists of a set of **vertices** and a set of **edges** joining pairs of vertices; no loops, multiple edges, or directed edges are allowed.



An **induced subgraph** of a graph consists of a subset of the vertex set together with all edges contained in the subset. In other words we are not allowed to delete edges within our chosen vertex set.

Rado's universal graph



In 1964, Richard Rado published a construction of a countable graph which was **universal**. This means that every finite or countable graph occurs as an induced subgraph of Rado's graph.

Rado's construction

The vertex set of Rado's graph R is the set \mathbb{N} of natural numbers (including 0).

Given two vertices x and y , with $x < y$, we join x to y if, when y is written in base 2, its x -th digit is 1 – in other words, if we write y as a sum of distinct powers of 2, one of these powers is 2^x .

Don't forget that the graph is undirected! Thus

- ▶ 0 is joined to all odd numbers;
- ▶ 1 is joined to 0 and to all numbers congruent to 2 or 3 (mod 4).
- ▶ ...

Problem

Does R have any non-trivial symmetry? And why is this very special graph the most famous infinite graph?

The random graph



Meanwhile, Rado's fellow Hungarians Paul Erdős and Alfred Rényi showed the following theorem:

Theorem

There is a countable graph R with the following property: if a random graph X on a fixed countable vertex set is chosen by selecting edges independently at random with probability $\frac{1}{2}$, then the probability that X is isomorphic to R is equal to 1.

The proof

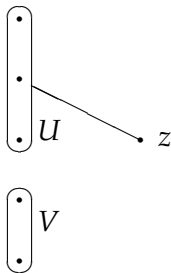
I will show you the proof.

I claim that one of the distinguishing features of mathematics is that you can be convinced of such an outrageous claim by some simple reasoning. I do not believe this could happen in any other subject.

Property (*)

The proof depends on the following property, which a graph may or may not possess:

- (*) Given two finite disjoint sets U and V of vertices, there is a vertex z which is joined to every vertex in U and to no vertex in V .



The point z is called a **witness** for the sets U and V .

Outline of the proof

I will prove:

Fact 1. With probability 1, a random countable graph satisfies (*).

Fact 2. Any two countable graphs satisfying (*) are isomorphic.

Then you will be convinced!

Proof of Fact 1

We use from measure theory the fact that a countable union of null sets is null. We are trying to show that a countable graph fails (*) with probability 0; since there are only countably many choices for the (finite disjoint) sets U and V , it suffices to show that for a fixed choice of U and V the probability that no witness z exists is 0.

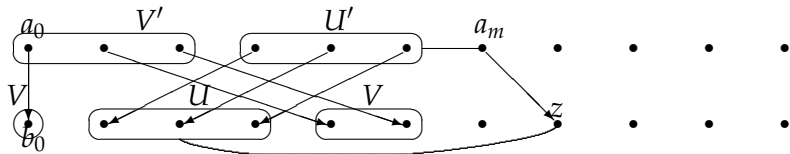
Suppose that $|U \cup V| = n$. Then the probability that a given vertex z is not the required witness is $1 - \frac{1}{2^n}$.

Since all edges are independent, the probability that none of z_1, z_2, \dots, z_N is the required witness is $(1 - \frac{1}{2^n})^N$, which tends to 0 as $N \rightarrow \infty$.

So the event that no witness exists has probability 0, as required.

Proof of Fact 2

We use a method known to logicians as “back and forth”. Suppose that Γ_1 and Γ_2 are countable graphs satisfying (*): enumerate their vertex sets as (a_0, a_1, \dots) and (b_0, b_1, \dots) . We build an isomorphism ϕ between them in stages.



At stage 0, map a_0 to b_0 .

At even-numbered stages, let a_m the first unmapped a_i . Let U' and V' be its neighbours and non-neighbours among the vertices already mapped, and let U and V be their images under ϕ . Use (*) in graph Γ_2 to find a witness v for U and V . Then map a_m to z .

Fact 2, continued

At odd-numbered stages, go in the other direction, using (*) in Γ_1 to choose a pre-image of the first unmapped vertex in Γ_2 . This approach guarantees that every vertex of Γ_1 occurs in the domain, and every vertex of Γ_2 in the range, of ϕ ; so we have constructed an isomorphism.

The proof is finished. This is a fine example of a non-constructive existence proof: if almost all graphs have the property, then certainly a graph with the property exists. Erdős and Rényi didn't bother with an explicit construction.

Had we only gone "forward", we would only use property (*) in Γ_2 , and we would have constructed an embedding, but could not guarantee that it is onto.

The back-and-forth method is often credited to Georg Cantor, but it seems that he never used it, and it was invented later by E. V. Huntington.

Properties of R

Recall that a countable graph Γ is **universal** if every finite or countable graph can be embedded into Γ as induced subgraph.

Fact 3. R is universal (for finite and countable graphs).

To see this, revisit the back-and-forth “machine” but use it only in the forward direction. As we saw, this only requires $(*)$ to hold in Γ_2 , and delivers an embedding of Γ_1 in Γ_2 .

A graph Γ is **homogeneous** if every isomorphism between finite induced subgraphs of Γ can be extended to an automorphism of Γ . (This is a very strong symmetry condition.)





Fact 4. R is homogeneous.

To see this, take $\Gamma_1 = \Gamma_2 = R$, and start the back-and-forth machine from the given finite isomorphism.

Randomness and symmetry

The fact that the random graph is highly symmetric is surprising, for several reasons.

First, for finite graphs, the more symmetric a graph, the smaller its probability of occurrence:

Graph				
Symmetries	6	2	2	6
Probability	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

In fact, the probability of any non-trivial symmetry tends rapidly to 0 as the number of vertices increases.

Indeed, the theorem of Erdős and Rényi was a short appendix to a long paper showing that most *finite* graphs are “as far from symmetry” as possible.

Second, recall the definition of Rado's graph R :

- ▶ Vertex set \mathbb{N}
- ▶ For $x < y$, x and y joined if the x -th binary digit of y is 1.

I mentioned the problem of finding a non-trivial symmetry of this graph. There seems to be no simple formula for one!

Rado's graph is indeed an example of the random graph. To prove this, all we have to do is to verify condition (*). This is a straightforward exercise.

A number-theoretic construction

Since the prime numbers are “random”, we should be able to use them to construct the random graph. Here’s how.

Recall that, if p is an odd prime not dividing a , then a is a **quadratic residue** (mod p) if the congruence $a \equiv x^2 \pmod{p}$ has a solution, and a **quadratic non-residue** otherwise. A special case of the law of **quadratic reciprocity**, due to Gauss, asserts that if the primes p and q are congruent to 1 (mod 4), then p is a quadratic residue (mod q) if and only if q is a quadratic residue (mod p).

So we can construct a graph whose vertices are all the prime numbers congruent to 1 (mod 4), with p and q joined if and only if p is a quadratic residue (mod q): the law of quadratic reciprocity guarantees that the edges are undirected.

This graph is isomorphic to the random graph!

To show this we have to verify condition (*). So let U and V be finite disjoint sets of primes congruent to 1 (mod 4). For each $u_i \in U$ let a_i be a fixed quadratic residue (mod u_i); for each $v_j \in V$, let b_j be a fixed quadratic non-residue mod v_j .

By the Chinese Remainder Theorem, the simultaneous congruences

- ▶ $z \equiv a_i \pmod{u_i}$ for all $u_i \in U$,
- ▶ $z \equiv b_j \pmod{v_j}$ for all $v_j \in V$,
- ▶ $z \equiv 1 \pmod{4}$,

have a solution modulo $4 \prod u_i \prod v_j$. By Dirichlet's Theorem, this congruence class contains a prime, which is the required witness.

The Skolem paradox

The **downward Löwenheim–Skolem theorem** of model theory says that a consistent theory in a countable first-order language has a countable model.

The **Skolem paradox** is this: There is a theorem of set theory (for example, as axiomatised by the Zermelo–Fraenkel axioms) which asserts the existence of uncountable sets. Assuming that ZF is consistent (as we all believe!), how can this theory have a countable model?

My point here is not to resolve this paradox, but to use it constructively.

A set-theoretic construction

Let M be a countable model of the Zermelo–Fraenkel axioms for set theory. Then M consists of a collection of things called “sets”, with a single binary relation \in , the “membership relation”.

Form a graph on the set M by joining x and y if either $x \in y$ or $y \in x$.

This graph turns out to be the random graph!

Indeed, the precise form of the axioms is not so important. We need a few basic axioms (Empty Set, Pairing, Union) and, crucially, the Axiom of Foundation, and that is all. It does not matter, for example, whether or not the Axiom of Choice holds.

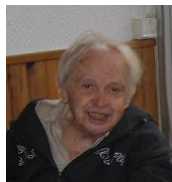
Back to Rado's graph

In the set-theoretic construction, it doesn't matter whether the axiom of infinity holds or not.

There is a simple description of a model of set theory in which the negation of the axiom of infinity holds (called **hereditarily finite set theory**). We represent sets by natural numbers. We encode a finite set $\{a_1, \dots, a_r\}$ of natural numbers by the natural number $2^{a_1} + \dots + 2^{a_r}$. (So, for example, 0 encodes the empty set.)

When we apply the construction of "symmetrising the membership relation" to this model, we obtain Rado's description of his graph!

Rolling back the years, 1



In fact, fifteen years earlier, Roland Fraïssé had asked the question: which homogeneous relational structures exist? Fraïssé defined the **age** of a relational structure M to be the class $\text{Age}(M)$ of all finite structures embeddable in M (as induced substructure). In terms of this notion, he gave a necessary and sufficient condition.

Fraïssé classes and Fraïssé limits

Theorem

A class \mathcal{C} of finite structures is the age of a countable homogeneous relational structure M if and only if

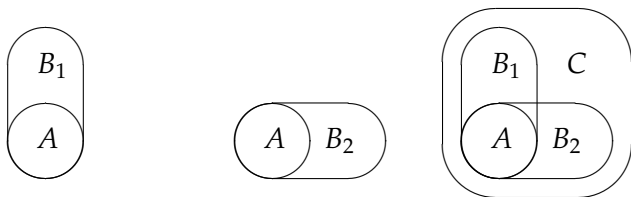
- ▶ *\mathcal{C} is closed under isomorphism;*
- ▶ *\mathcal{C} is closed under taking induced substructures;*
- ▶ *\mathcal{C} contains only countably many non-isomorphic structures;*
- ▶ *\mathcal{C} has the **amalgamation property** (see next slide).*

If these conditions hold, then M is unique up to isomorphism.

A class \mathcal{C} satisfying these conditions is a **Fraïssé class**, and the countable homogeneous structure M is its **Fraïssé limit**.

The amalgamation property

The amalgamation property says that two structures B_1, B_2 in the class \mathcal{C} which have substructures isomorphic to A can be “glued together” along A inside a structure $C \in \mathcal{C}$:



Note that the intersection of B_1 and B_2 may be larger than A .

Examples

Each of the following classes is a Fraïssé class; the proofs are exercises. Thus the corresponding universal homogeneous Fraïssé limits exist.

Fraïssé class	Fraïssé limit
Graphs	Rado's graph
Triangle-free graphs	Henson's graph
Graphs with bipartition	Generic bipartite graph
Total orders	$(\mathbb{Q}, <)$
Partial orders	Generic poset
Permutation patterns	Generic biorder

There are *many* others!

Rolling back further



A quarter of a century earlier, these ideas had already been used by the Soviet mathematician P. S. Urysohn. He visited western Europe with Aleksandrov, talked to Hilbert, Hausdorff and Brouwer, and was drowned while swimming in the sea at Batz-sur-Mer in south-west France at the age of 26 in 1924. Among his many contributions to topology was the theorem discussed below. The paper was completed from Urysohn's unpublished work by Aleksandrov and published in 1926.

Urysohn's theorem

A **Polish space** is a metric space which is **complete** (Cauchy sequences converge) and **separable** (there is a countable dense set). A metric space M is **homogeneous** if any isometry between finite subspaces extends to an isometry of M .

Theorem

There exists a homogeneous Polish space containing a copy of every finite metric space, and it is unique up to isometry.

This unique metric space is known as the **Urysohn space**. Its study has been popularised in recent years by Anatoly Vershik.

The proof

Here, in modern terminology, is what Urysohn did.

We cannot apply Fraïssé's Theorem directly to obtain this result, since there are uncountably many 2-element metric spaces up to isometry (one for each positive real number).

Instead, use the class of finite **rational** metric spaces (those with all distances rational). This is a Fraïssé class, whose Fraïssé limit is a countable universal homogeneous rational metric space.

Its completion is easily seen to be the required Polish space.

Related constructions

Various other types of metric spaces form Fraïssé classes. These include

- ▶ The class of **integral** metric spaces, those with all distances integers. The Fraïssé limit is a kind of universal distance-transitive graph.
- ▶ The class of metric spaces with all distances 1 or 2. The Fraïssé limit is the **random graph**!

Let M be the Fraïssé limit of the class of metric spaces with all distances 1 or 2; form a graph by joining two points if their distance is 1. Since the graph is homogeneous, if v and w are two vertices at distance 2, there is a vertex at distance 1 from both. Thus the distance in M coincides with the graph distance in this graph. The graph is universal and homogeneous, and so is R .

Reversing the arrows

A category theorist, looking at Fraïssé's construction, would draw a commutative diagram, with arrows representing embeddings.

She would then ask for a similar theorem, with the arrows reversed, and representing projections.

There is such a theorem, the **projective Fraïssé theorem**. Rather than describe it in detail, I will give you one application.

Cantor space

The famous **Cantor set** is given by the **middle third** construction, starting with the unit interval, and successively removing the middle third from each interval.

It consists of the real numbers in the unit interval whose base 3 representation involves 0s and 2s only.

Topologically, it is homeomorphic to a product of countably many copies of a 2-element discrete space, and hence it is compact and totally disconnected.

If we replace 0 and 2 in base 3 by 0 and 1 in base 2, we make a huge difference to the topology: we now get the unit interval, which is connected. This change occurs because a few real numbers have two base 2 representations, and so the map from Cantor space to the unit interval is not quite one-to-one.

The pseudo-arc

What does a typical closed connected subset of the unit square look like?

We have to be careful about the word “typical”. In a probability space this can mean “a set of measure 1”, but here we don’t have a measure. Instead we use a notion from **Baire category**: in a complete metric space, a set is **residual** if it contains a countable intersection of open dense subsets. Residual sets behave like complements of null sets: they are non-empty, meet every open set, and any two (or countably many) of them intersect in a residual set.

The metric we use on closed subsets of the square is **Hausdorff metric**: two sets are within distance ϵ if every point of one is within distance ϵ from some point of the other.

Now it turns out that there is a space \mathbb{P} such that, in the set of closed connected subsets of the unit square with the Hausdorff metric, the elements homeomorphic to \mathbb{P} form a residual set.



The space \mathbb{P} is the **pseudo-arc**.

Several different constructions were given (the first by Knaster in 1922), but R. H. Bing showed that they all produced the same object, and that the homeomorphism group of \mathbb{P} acts transitively on its points.

Moreover, the statement in the first paragraph remains true if we replace the unit square by the unit hypercube in \mathbb{R}^n for any $n \geq 2$, or in Hilbert space.

Its topological definition might suggest that it cannot be constructed by discrete methods, but this is not so . . .

... as projective Fraïssé limit (almost)

Consider the class \mathcal{P} of **reflexive paths**, graphs which consist of a finite path with a loop at each vertex. Irwin and Solecki show that \mathcal{P} is a projective Fraïssé class, so has a projective Fraïssé limit P .

Thus P has the structure of a graph and the topology of the Cantor set.

They show further that the graph structure on P consists of isolated vertices and edges (with loops) only; thus, an equivalence relation with all equivalence classes of size 1 or 2. Taking the quotient of P by this equivalence relation gives the **pseudo-arc** \mathbb{P} .

The last step mirrors the step from Cantor space to the unit interval.

Using this, Solecki and Tsankov were able to give a new proof of Bing's theorem that the pseudo-arc has a transitive homeomorphism group.