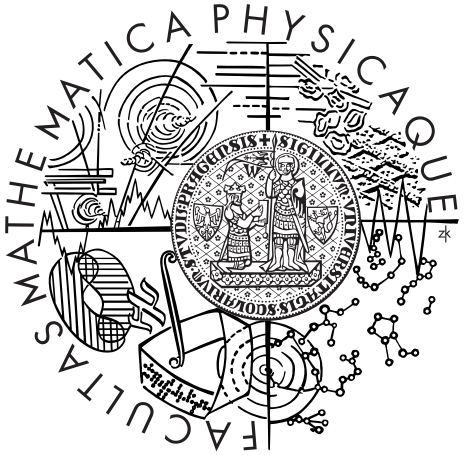
Charles University in Prague
Faculty of Mathematics and Physics

**DOCTORAL THESIS**




David Hartman

# Extension properties of structures

Computer Science Institute of Charles University

Supervisor of the doctoral thesis: Prof. RNDr. Jaroslav Nešetřil, DrSc.
Study programme: Informatics
Specialization: 414

Prague 2014

I would like to express my gratitude to my advisor Prof. RNDr. Jaroslav Nešetřil, DrSc. for patient leadership and the inspiration he gave me to explore many areas of mathematics as well as his inspiration in personal life more generally. Furthermore, I would like to thank my coauthors Jan Hubička and Dragan Mašulović for helping me to establish ideas and for the realization of these ideas in their final published form. Moreover thanks are due to Peter J. Cameron for posing several questions about classes of homomorphism-homogeneous structures in our personal communication that subsequently led to one of the publications as well as to John K. Truss whose interest in my work after the first Workshop on Homogeneous Structures supported my further progress. Many thanks belong to Andrew Goodall who provide an extensive feedback for many parts of the thesis.

Finally, the greatest thanks belong to my family, my beloved wife Mirečka and our children Emma Alexandra and Hugo, as altogether the main source of joy, happiness and in fact meaning of life.

I declare that I carried out this doctoral thesis independently, and only with the cited sources, literature and other professional sources.

I understand that my work relates to the rights and obligations under the Act No. 121/2000 Coll., the Copyright Act, as amended, in particular the fact that the Charles University in Prague has the right to conclude a license agreement on the use of this work as a school work pursuant to Section 60 paragraph 1 of the Copyright Act.

In ..... date .....  signature of the author

**Název práce:** Rozšiřující vlastnosti struktur
**Autor:** David Hartman
**Katedra:** Informatický ústav Univerzity Karlovy
**Vedoucí disertační práce:** Prof. RNDr. Jaroslav Nešetřil, DrSc.
**Abstrakt:** Tato práce rozebírá vlastnost relačních struktur, která implikuje jejich vysokou symetričnost. Strukturu nazveme homogenní pokud lze libovolné lokální zobrazení rozšířit na zobrazení nad celou strukturou a to pro libovolnou volbu konečné vzorové množiny. Typ lokálního a globálního zobrazení potom určuje různé typy homogenity. Prominentní místo má ultrahomogenita, která označuje strukturu, pro kterou libovolný lokální isomorfismus nad konečnými podstrukturami je rozšiřitelný na automorfismus. Na rozdíl od grafů je klasifikace ultrahomogenních relačních struktur stále otevřeným problémem. Cílem práce je charakterizovat "vzdálenost" od homogenity a to dvěma způsoby. Nejprve zvyšuje "složitost struktury" přidáváním relací a sleduje změny klasifikace homogenních struktur. To vede k několika klasifikacím homomorfně-homogenních $L$-obarvitelných grafů pro různé $L$, kde $L$-obarvitelný graf je graf, kde vrcholy a hrany dostávají množiny barev z částečně uspořádané množiny $L$. Na to navazují výsledky a diskuze nad hierarchií tříd definovanou skrze různé typy homogenity s ohledem na koincidenci jednotlivých tříd. Druhý pohled zkoumá pro dané struktury jak minimálně rozšířit jejich jazyk, abychom dosáhli homogenity. Výsledky se týkají relační komplexity konečných grafů a dále její meze pro nekonečně spočetné relační struktury definované třídou zakázaných homomorfismů.
**Klíčová slova:** relační struktury, ultrahomogení struktury, homomorfně homogenní struktury, obarvené grafy, relační složitost, komplexní sítě

**Title:** Extension property of structures
**Author:** David Hartman
**Department:** Computer Science Institute of Charles University
**Supervisor:** Prof. RNDr. Jaroslav Nešetřil, DrSc.
**Abstract:** This work analyses properties of relational structures that imply a high degree of symmetry. A structure is called homogeneous if every mapping from any finite substructure can be extended to a mapping over the whole structure. The various types of these mappings determine corresponding types of homogeneity. A prominent position belongs to ultrahomogeneity, for which every local isomorphism can be extended to an automorphism. In contrast to graphs, the classification of ultrahomogeneous relational structures is still an open problem. The task of this work is to characterize "the distance" to homogeneity using two approaches. Firstly, the classification of homogeneous structures is studied when the "complexity" of a structure is increased by introducing more relations. This leads to various classifications of homomorphism-homogeneous $L$-colored graphs for different $L$, where $L$-colored graphs are graphs having sets of colors from a partially ordered set $L$ assigned to vertices and edges. Moreover a hierarchy of classes of homogeneous structures defined via types of homogeneity is studied from the viewpoint of classes coincidence. The second approach analyses for fixed classes of structures the least way to extend their language so as to achieve homogeneity. We obtain results about relational complexity for finite graphs and bounds on this complexity for countably infinite structures defined via classes of forbidden homomorphisms.
**Keywords:** relational structures, ultrahomogeneous structures, homomorphism-homogeneous structures, colored graphs, relational complexity, complex networks

# Contents

# 1. Introduction

> The mathematical sciences particularly exhibit order, symmetry, and limitation; and these are the greatest forms of the beautiful.
>
> Aristotle

There is a long-standing effort, albeit not precisely defined, to discover and explore beauty in mathematical objects of all kinds. The roots of beauty are often found to be connected with symmetries of the corresponding object. There are many ways how symmetry can be defined, usually through different representations of the object under study. Using notions from model theory, an interesting type of symmetry – and from a specific point of view one of the most fundamental – is achieved when mathematical objects are studied from the perspective of "structure-preserving" extensions of morphisms over the structure itself. To be able to define this kind of symmetry in a more formal way, let us first define the necessary notions that are required.

A *relational structure* $\mathbf{A}$ is a pair $(A, R_{\mathbf{A}})$, where $R_{\mathbf{A}}$ is a tuple $(R^i_{\mathbf{A}} : i \in I)$ of relations such that $R^i_{\mathbf{A}} \subseteq A^{\delta_i}$ (i.e. $R^i_{\mathbf{A}}$ is a $\delta_i$-ary relation on $A$). It should be mentioned that constant and function symbols can also be defined for general structures, but we do not consider these here – for more information see [67]. The underlying set $A$ is called the *domain* of $\mathbf{A}$, whose cardinality is the corresponding cardinality of the whole relational structure, i.e. $|\mathbf{A}| = |A|$. Elements of the domain are usually denoted by $a_1, a_2, \ldots$, or by $x_1, x_2, \ldots$. A finite sequence of elements $(a_1, a_2, \ldots, a_n)$ or $(x_1, x_2, \ldots, x_n)$ is called an *n-tuple* or simply a *tuple*, and denoted by $\overline{a}$ or $\overline{x}$. The *signature* $L$ of a relational structure $\mathbf{A}$ is a set of $n$-ary relational symbols given for each separate $n$. A signature represents a language from which the relational structure is constructed and it can be usually assumed to be read simply from the definition of the structure. When there is a need to specify the language for a particular structure we call such a structure an $L$-structure. The class of all (countable) relational structures with signature $L$ is denoted by $\mathrm{Rel}(L)$.

One of the structures that features prominently in the sequel is that of an undirected graph. An *undirected graph*, or simply a *graph*, is a relational structure with domain $V$ whose elements are called *vertices* and one binary irreflexive symmetric relation $E$. The set of unordered pairs $\{u, v\}$ such that $E(u, v)$ are are called *edges*. Vertices are usually denoted by $v_1, v_2, \ldots$ or $u_1, u_2, \ldots$. Edges will be denoted by $e_1, e_2, \ldots$, or simply by concatenating vertex symbols, so that for example $v_iv_j$ denotes the edge joining vertices $v_i$ and $v_j$. If the undirected nature of an edge has to be stressed we write $\{v_i, v_j\}$ rather than $v_iv_j$. This kind of structure is usually abbreviated as $G = (V, E)$. For such a graph *loops*, i.e. edges whose endpoints are the same vertex, are not allowed, which follows from the irreflexive property assumed for the edge relation. A *graph with loops allowed* is a graph where the corresponding edge relation is not necessarily irreflexive. The vertex set of a given graph $G$ is denoted by $V(G)$, and similarly its edge set is denoted by $E(G)$.

The source of symmetry within these relational structures lies in a specific handling of various mappings defined over them. For structures $\mathbf{A} = (A, R_{\mathbf{A}})$ and $\mathbf{B} = (B, R_{\mathbf{B}})$ a *homomorphism* $f : \mathbf{A} \to \mathbf{B}$ is a mapping $f : A \to B$ such that $(x_1, x_2, \ldots, x_{\delta_i}) \in R^i_{\mathbf{A}}$ implies $(f(x_1), f(x_2), \ldots, f(x_{\delta_i})) \in R^i_{\mathbf{B}}$ for each $i \in I$. If $f$ is further one-to-one then $f$ is called a *monomorphism*. An *isomorphism* $g : \mathbf{A} \to \mathbf{B}$ is a bijective mapping $g : A \to B$ such that $(x_1, x_2, \ldots, x_{\delta_i}) \in R^i_{\mathbf{A}} \Leftrightarrow (g(x_1), g(x_2), \ldots, g(x_{\delta_i})) \in R^i_{\mathbf{B}}$ for each $i \in I$. An isomorphism from a structure to itself is called an *automorphism*. On the other hand, if the structure is isomorphically mapped into, but not onto, a different one then it is called an embedding. Formally an *embedding* $h : \mathbf{A} \to \mathbf{B}$ is a injective mapping $h : A \to B$ such that for each $i \in I$ and each $\delta_i$-tuple $(x_1, x_2, \ldots, x_{\delta_i}) \in A$ it holds that $(x_1, x_2, \ldots, x_{\delta_i}) \in R^i_{\mathbf{A}} \Leftrightarrow (h(x_1), h(x_2), \ldots, h(x_{\delta_i})) \in R^i_{\mathbf{B}}$. Similarly, an *endomorphism* is a homomorphism from a structure to itself. The set of all automorphisms of a structure $\mathbf{A}$ form its *automorphism group* $\mathrm{Aut}(\mathbf{A})$.

Let $f$ be a mapping from $\{x_1, x_2, \ldots, x_n\}$ to $\{y_1, y_2, \ldots, y_n\}$ such that $f(x_i) = y_i$. For such a mapping the following shorthand notation is used throughout the paper:

$$f = \begin{pmatrix} x_1 & x_2 & \ldots & x_n \\ y_1 & y_2 & \ldots & y_n \end{pmatrix}.$$

We call $\mathbf{B}$ a *substructure* of $\mathbf{A}$, in symbols $\mathbf{B} \subseteq \mathbf{A}$, when $B \subseteq A$ and the inclusion map $f : B \to A$ is an embedding. We ought to mention here the following fact. Let $\mathbf{A}$ be an $L$-structure and $X$ the set of its elements. Using [67, Lemma 1.2.2] it can be easily shown that there is a unique smallest substructure $\mathbf{B} \subseteq \mathbf{A}$ such that $X \subseteq \mathbf{B}$. While we are working only with relational structures, any structure generated in this way from finite $X$ is also finite. See [67, ex. 6, p.10].

Similarly to relational structures we can define for a given graph $G$ a *subgraph* $H$ as a graph on vertex set $V(H) \subseteq V(G)$ having $E(H) \subseteq E(G)$. Note that this does not correspond with relational structures completely, since we can omit some edges from consideration while in the case of structures these are induced automatically from the embedding. For these reasons we can define for a graph $G$ an *induced subgraph* $H'$ that has $V(H') \subseteq V(G)$ and that is an embedding of $H'$ into $G$. The condition for edges can be equivalently defined as all edges of $G$ induced by set $V(H')$, i.e. $E(H') = \binom{V(G)}{2} \cap E(G)$. The notation $\binom{V(G)}{2}$ for the set of all subsets of $V(G)$ of size 2 will be frequently used in the sequel.

This is an introductory chapter whose purpose is to familiarize the reader with notions that are used further in the text. It therefore contains virtually nothing new and most if not all proofs are adapted from the referenced sources, only adjusted accordingly so as to fit the whole story and notation used here.

## 1.1 Symmetry of structures

As stated above, the kind of symmetry that this work is interested in deals with various kinds of morphisms defined on a structure. Roughly speaking, symmetry of this kind is defined as the ability to extend any existing local morphism of the structure of an agreed type to a morphism over the whole structure, again of an agreed (but not necessarily the same) type. One immediate example of

such a property is that of vertex transitivity. A *vertex transitive graph* is a graph $G$ where for all pairs of vertices $x$ and $y$ there exists an automorphism $f \in \mathrm{Aut}(G)$ such that $y = f(x)$. This means roughly that all the vertices of a vertex transitive graph look the same in the sense of their position in the graph. There are several obvious examples of vertex transitive graphs like complete graphs or cycles. Another well known example of a graph of this type is the *Petersen graph* shown in Fig. 1.1.

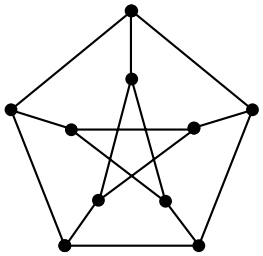


Figure 1.1: Petersen graph is vertex transitive.

It is possible to put the condition of transitivity on edges rather than vertices, which yields an *edge transitive* graph, i.e. a graph where for every pair of edges there exists an automorphism that maps one to the other. All the above-mentioned graphs are also edge transitive. The complete bipartite graph $K_{m,n}$ where $m \neq n$ is edge transitive but not vertex transitive. Both vertex and edge transitivity have been studied for a long time and are still a subject of further research – see for example [89]. A central theme of this work is to study symmetries that generalize these transitivity properties.

## 1.2 Ultrahomogeneity

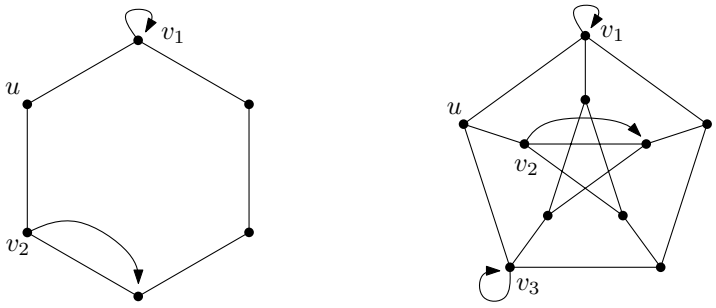
For a general relational structure we consider the following symmetry property.

**Definition 1.2.1.** A relational structure $\mathbf{A}$ is called *ultrahomogeneous* if every isomorphism between two induced finite substructures of $\mathbf{A}$ can be extended to an automorphism of $\mathbf{A}$.

Ultrahomogeneity describes a high degree of symmetry for relational structures. Staying in the class of graphs, we can mention complete graphs as an immediate example. On the other hand, cycles, which are vertex and edge transitive, are ultrahomogeneous only up to size 5 – see section 1.4. To see that the cycle of length 6 is not ultrahomogeneous one can simply map two vertices with a common neighbor to another pair of vertices without such a common neighboring vertex, which results in the inability of any automorphism to find an image for this common neighbor, as shown in Fig. 1.2. Similarly for the Petersen graph one can use an independent set of size 3 to show that it is not ultrahomogeneous.

### 1.2.1 The Rado graph

For infinite graphs the situation is also interesting. Omitting obvious examples like infinite complete graphs we mention what is known as the Rado graph. This




(a) $C_6$ is not ultrahomogeneous (b) Petersen is not ultrahomogeneous

Figure 1.2: Examples of non-ultrahomogeneous finite graphs, where an independent set composed of vertices $v_i$ is mapped by a local isomorphism for which it is impossible to to find an image for a common neighbor of this set.

graph is defined by a famous theorem of Erdős and Renyi.

**Theorem 1.2.1** (Erdős and Renyi [40])**.** *There is a countable graph $R$ with the property that a random countable graph (edges chosen independently with probability 1/2) is almost surely isomorphic to $R$.*

When discussing homogeneity of this graph notation and ideas are used from [16], which is an excellent source of information on the countable random graph. To show that this graph is ultrahomogeneous one can make use of the following graph property.

**Definition 1.2.2.** A graph $G$ has property * if for any finite disjoint sets $U$ and $V$ of vertices, there exists a vertex $z$ joined to every vertex $u_i \in U$ and to none of the vertices $v_j \in V$.

A vertex $z$ satisfying the given property is called *correctly joined*. Obviously, a graph satisfying this property is infinite. On the other hand finding an infinite graph meeting this requirement is not immediate. However one can show that for the infinite random graph the property * holds – see below. This was already shown by Erdős and Renyi [40]. The following formulation of the result is taken from [16].

**Proposition 1.2.2.** *With probability* 1*, a countable random graph satisfies property *.*

*Proof.* The statement of the proposition can be proved by showing that property * fails with probability 0. In other words, the set of graphs for which * fails for some given set of vertices $U = \{u_1, u_2, \ldots, u_m\}$ and $V = \{v_1, v_2, \ldots, v_n\}$ is a null set. There are only countably many $m$ and $n$ and consequently countably many choices of $U$ and $V$ so that it suffices to prove the previously mentioned claim for a fixed choice of $U$ and $V$, the union of countably many null sets being a null set.

The remaining step is to calculate the probability of this event, namely the probability that for any $z_1, z_2, \ldots, z_N$ distinct from $u_1, u_2, \ldots, u_m$ and $v_1, v_2, \ldots, v_n$ it is the case that all the $z_i$ are not correctly joined. For a particular $z_i$ this can be expressed as:

$$1 - \frac{1}{2^{m+n}}$$

and since the corresponding events are independent the resulting formula for probability that none of $z_1, z_2, \ldots, z_N$ are correctly joined is

$$\left(1 - \frac{1}{2^{m+n}}\right)^N$$

Finally, this is easily seen to tend to 0 as $N \to \infty$. □

Using the property * of the countable random graph $R$ one can show that the Rado graph is in fact ultrahomogeneous.

**Proposition 1.2.3.** *The countable random graph $R$ is ultrahomogeneous.*

*Proof.* The countable random graph $R$ satisfies property * by Proposition 1.2.2. Now let $f$ be any isomorphism from a finite set of vertices $X = \{x_1, x_2, \ldots, x_n\}$ to $R$ and let $x_{n+1} \in V(R)$ be another vertex outside $X$. The task is to show that $f$ can be extended so that its domain is $X \cup \{x_{n+1}\}$. Let $U$ be the set of neighbors of $x_{n+1}$ in $X$ and $V = X \setminus U$ the set of non-neighbors of $x_{n+1}$ in $X$. Now $f(x_{n+1})$ has to be connected to all the vertices from $f(U)$ and none of the vertices from $f(V)$. Using property * we know that such a vertex exists, so we can extend $f$ by mapping $x_{n+1}$ to this vertex. In this way any given local isomorphism can be extended by one vertex.

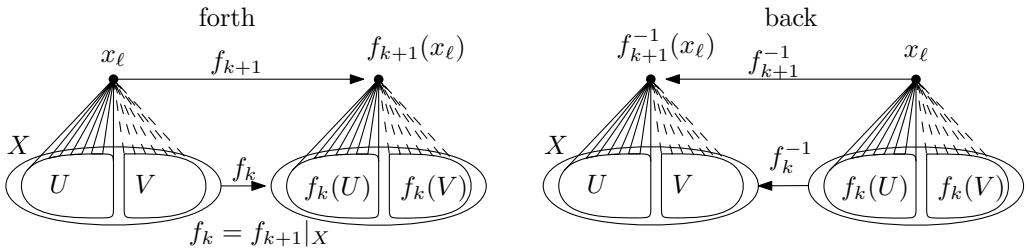



Figure 1.3: Extending local isomorphism $f_k$ using the back-and-forth construction. Solid lines represents edges and dashed lines non-edges.

In the second part of the proof the previously mentioned process is used to construct an automorphism that extends the given local isomorphism $f$ by using a back-and-forth argument [75, 20]. Enumerate the vertices of $R$ as $\{x_1, x_2, \ldots\}$. Assume that $f_0 = \emptyset$ is a finite isomorphism. Let us construct finite isomorphisms $f_k$, $k \in \mathbb{N}$, as follows. Suppose that $f_k$ has been constructed. The next step depends on the value of $k$, where even values are used for the forward construction and odd values for the backward. If $k$ is even, extend $f_k$ to $f_{k+1}$ by adding a vertex $x_\ell$ to the domain of $f_k$, where $\ell = \min\{i : x_i \in V(R) \text{ and } x_i \notin \text{domain of } f\}$. On the other hand, if $k$ is odd then extend $f_k$ to $f_{k+1}$ by adding a vertex $x_\ell$ to the range of $f_k$, where $\ell = \min\{j : x_j \in V(R) \text{ and } x_j \notin \text{range of } f_k\}$. Finally, to obtain the required isomorphism, take $f$ to be the union of all these partial isomorphisms, where the back-and-forth construction ensures that all the vertices of $R$ are in both the range and the domain of $f$. □

To explore different structures than graphs, let us consider partial orders, where an interesting example of an ultrahomogeneous structure is $(\mathbb{Q}, <)$, where $<$ stands for usual order on rationals. The homogeneity of this structure can be for example shown by applying the same back-and-forth mechanism described in the proof of Proposition 1.2.3 for the case of the countable random graph $R$.

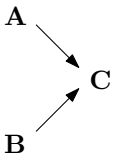
### 1.2.2 The Fraïssé method

Ultrahomogeneity is tightly connected with results of Roland Fraïssé, who provided a simple construction of countable ultrahomogeneous structures. Following his work, we can define the *age* of relational structure $\mathbf{A}$, denoted by $\operatorname{Age}\mathbf{A}$, as the class of all finite structures that can be embedded in $\mathbf{A}$. Since we are only interested in structures up to isomorphism we can consider the age of $\mathbf{A}$ to be the set of all finite substructures of $\mathbf{A}$. We say that structure $\mathbf{A}$ is *younger* than structure $\mathbf{B}$ when $\operatorname{Age}(\mathbf{A}) \subseteq \operatorname{Age}(\mathbf{B})$. To describe the age of a relational structure more deeply it is valuable to consider the following properties of a set of finite relational structures.

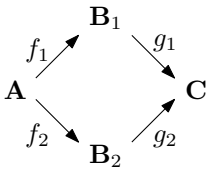
**Definition 1.2.3.** A class of finite relational structures $\mathcal{K}$ has

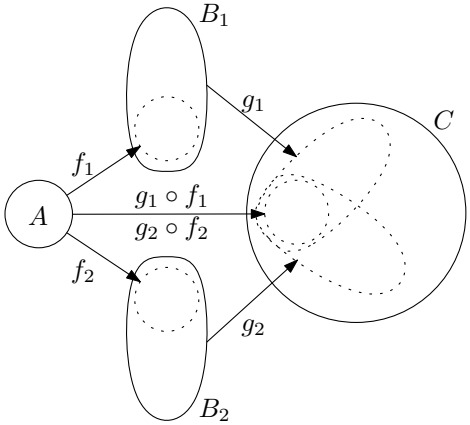
1. the **hereditary property** (abbreviated as **HP**) if for any $\mathbf{A} \in \mathcal{K}$ and $\mathbf{B}$ a finite substructure of $\mathbf{A}$ there exists $\mathbf{C} \in \mathbf{K}$ that is isomorphic to $\mathbf{B}$;

2. the **joint embedding property** (abbreviated as **JEP**) if for any $\mathbf{A}, \mathbf{B} \in \mathcal{K}$ there exists $\mathbf{C}$ such that both $\mathbf{A}$ and $\mathbf{B}$ are embeddable in $\mathbf{C}$;




3. the **amalgamation property** (abbreviated as **AP**) if for any $\mathbf{A}, \mathbf{B}_1, \mathbf{B}_2 \in \mathcal{K}$ and embeddings $f_i : \mathbf{A} \to \mathbf{B_i}$ (for $i = 1, 2$), there exist a relational structure $\mathbf{C}$ and embeddings $g_i : \mathbf{B_i} \to \mathbf{C}$ (for $i = 1, 2$) such that $g_1 \circ f_1 = g_2 \circ f_2$.

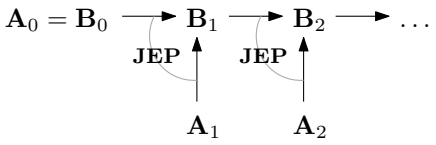



It is a simple observation that a class $\mathcal{K}$ that is the age of some relational structure has both **HP** and **JEP**. To prove the converse, i.e. that any class of finite structures having **HP** and **JEP** is the age of some relational structure, requires further machinery.

**Proposition 1.2.4** (Theorem 6.1.1 in [67])**.** *Let $\mathcal{K}$ be a finite or countable set of finite relational structures which has* **HP** *and* **JEP***. Then $\mathcal{K}$ is the age of some finite or countable relational structure.*

*Proof.* Let $(\mathbf{A}_i)_{i<\omega}$ be a list of all structures in $\mathcal{K}$. Define corresponding structures $\mathbf{B_i}$ as follows. First set $\mathbf{B_0} = \mathbf{A_0}$. Suppose $\mathbf{B_i}$ has already been defined. Then use **JEP** to find a structure $\mathbf{B_{i+1}}$ into which both $\mathbf{B_i}$ and $\mathbf{A_{i+1}}$ are embeddable.




Finally put $\mathbf{B} = \bigcup_{i<\omega} \mathbf{B_i}$. Following the construction, each $\mathbf{A_i}$ has an embedding into $\mathbf{B}$ and moreover $\mathbf{B}$ is at most countable. Therefore $\mathcal{K}$ is contained in the age of $\mathbf{B}$. Moreover any finite substructure of $\mathbf{B}$ is a finite substructure of some $\mathbf{B_i}$ and therefore of $\mathbf{B}$ by **HP**. □

The above statement with the previous observation provide a way to handle the ages of relational structures, however uniqueness of the resulting structure is still not ensured. For this to be the case, more properties have to be defined. One property that ensures uniqueness of the corresponding structure is the following.

**Definition 1.2.4.** A structure $\mathbf{D}$ has the *extension property* (abbreviated as **EP**) if for any $\mathbf{A}$ and $\mathbf{B}$ belonging to Age($\mathbf{D}$) such that $\mathbf{A} \subseteq \mathbf{B}$ and every embedding $f : \mathbf{A} \to \mathbf{D}$, there is an embedding $g : \mathbf{B} \to \mathbf{D}$ which extends $f$.

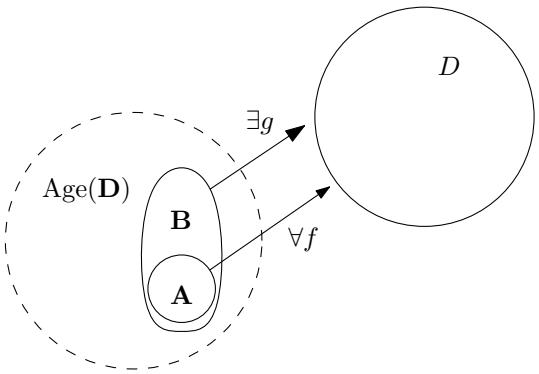



Figure 1.4: Structure $\mathbf{D}$ having the extension property.

One immediate observation is that ultrahomogeneous structures have the extension property. The other way round is not so direct.

**Lemma 1.2.5** (adapted from [67])**.** *Let* $\mathbf{D}$ *be a countable structure with the extension property. Then the following hold.*

1. *If there exists a countable structure* $\mathbf{C}$ *with the extension property having* Age($\mathbf{D}$) = Age($\mathbf{C}$) *then* $\mathbf{C}$ *is isomorphic to* $\mathbf{D}$*.*

2. $\mathbf{D}$ *is ultrahomogeneous.*

*Proof.* (sketch) To prove 1. we need to build an isomorphism between $\mathbf{C}$ and $\mathbf{D}$, which can be done using the extension property of both structures. Recalling ideas from the proof of Proposition 1.2.4, express both structures as the union of chains $(\mathbf{C_i})_{i<\omega}$ and $(\mathbf{D_i})_{i<\omega}$ of finite substructures and iteratively construct isomorphisms $f_i$ whose union is an isomorphism between $\mathbf{C}$ and $\mathbf{D}$.

Let $f_0 : \mathbf{C_0} \to \mathbf{D_0}$ be an isomorphism of finite substructures of $\mathbf{C}$ and $\mathbf{D}$, where $\mathbf{C_0}$ and $\mathbf{D_0}$ can be arbitrary substructures for which this holds. Furthermore let $f_i : \mathbf{C_i} \to \mathbf{D_i}$ be given. The next step depends on value of $i$. If $i$ is odd, then since $\mathrm{Age}(\mathbf{C}) = \mathrm{Age}(\mathbf{D})$ there is an isomorphism $g_{i+1} : \mathbf{C_{i+1}} \to \mathbf{D_{i+1}}$ and $\mathbf{D_{i+1}}$ is embeddable in $\mathbf{D}$ which means that there exists an embedding $h_{i+1} : \mathbf{D_{i+1}} \to \mathbf{D}$. Now $f_i \circ g_{i+1}^{-1}$ embeds $g_{i+1}(\mathbf{C_i})$ into $\mathbf{D}$. Moreover, using the extension property of $\mathbf{D}$, this embedding extends to an embedding $h_{i+1}$. Now let $f_{i+1} : \mathbf{C_{i+1}} \to \mathbf{D}$ be $h_{i+1} \circ g_{i+1}$, for which we have $f_i \subseteq f_{i+1}$. For even $i$ the process is similar but reverse using the fact that $f_{i+1}$ has $D_{i+1}$ in its image. This provides a chain of maps $(f_i)_{i<\omega}$ whose union is the desired isomorphism.

To prove 2. it is sufficient to use previous proof with $C = D$ and start with any predefined isomorphism $f_0$. □

Using the forth mechanism from the proof of Lemma 1.2.5 part 1 it is possible to show that into a structure $\mathbf{D}$ having **EP** one can embed any younger structure [67]. A structure having this property is called *universal*. Rather than going deep into this proof, which is similar to that given above, we shall show that the countable random graph $R$ is universal. This is actually one of the reasons why Rado was interested in this graph and why sometimes this graph is also called the Rado graph [131].

**Proposition 1.2.6** (Appeared in [16])**.** *Any finite or countable graph can be embedded as an induced subgraph of $R$.*

*Proof.* Let $G$ be any finite or countable graph. The task is to find an embedding of this graph into $R$. The construction is iterative, however unlike the proof of Proposition 1.2.3 only the forth step is used since property * (see Definition 1.2.2) cannot be generally assumed in $G$. Let us denote the set of vertices of $G$ by $X = \{x_1, x_2, \ldots\}$ and let $X_k = \{x_1, x_2, \ldots, x_k\}$. Assume further that there is a mapping $f_k : X \to R$ that is an isomorphism of the induced subgraph on $X_k$. The task is to find an image $z$ of the next vertex $x_{k+1} \in X$ such that extending the domain of $f_k$ to include the vertex $x_{k+1}$ again gives an isomorphism, call it $f_{k+1}$, this time on the induced substructure on vertex set $X_{k+1}$. The union of these partial isomorphisms $f = \bigcup f_k$ is the required embedding. The existence of the vertex $z$ is ensured by the * property of $R$. □

Proposition 1.2.6 has the interesting consequence that $R$ contains an infinite clique and an infinite independent set, both of which are further examples of ultrahomogeneous structures – their homogeneity is obvious. A general variant of Proposition 1.2.6 also shows that a countable structure possessing **EP** is universal for its age.

**Proposition 1.2.7** (Appeared in [19])**.** *An ultrahomogeneous structure is universal.*

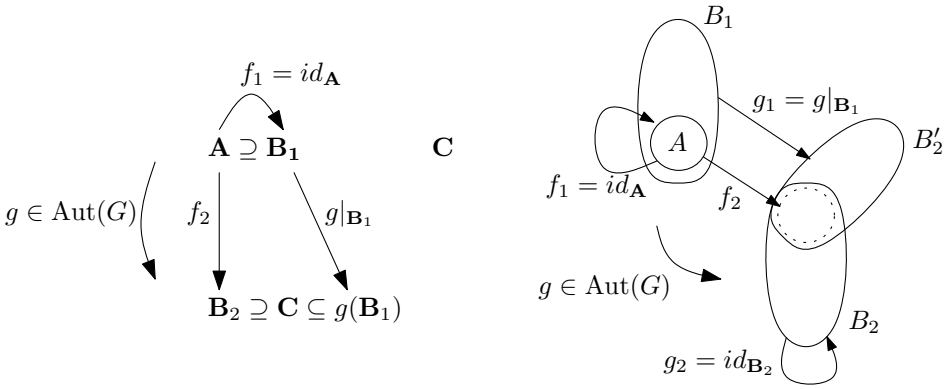
We shall now see what happens when we take the amalgamation property into consideration. The following famous theorem shows that a structure with corresponding age is ultrahomogeneous and moreover unique.

**Theorem 1.2.8** (Fraïssé theorem [44, 67])**.**

*(a) Let* $\mathbf{D}$ *be an ultrahomogeneous relational structure. Then* Age($\mathbf{D}$) *has the amalgamation property* **AP**.

*(b) Let* $\mathcal{K}$ *be a non-empty finite or countable set of finite relational structures which has* **HP**, **JEP** *and* **AP**. *Then there is an ultrahomogeneous relational structure* $\mathbf{D}$, *unique up to isomorphism, whose age is* $\mathcal{K}$.

*Proof.* (Sketch) For part (a) consider the relational structures $\mathbf{A}$, $\mathbf{B_1}$ and $\mathbf{B_2}$ and mappings $f_1$ and $f_2$ defined as for use in property **AP**. We may assume that both sets $\mathbf{B_i}$ for $i = 1, 2$ are substructures of $\mathbf{D}$. Without loss of generality we take the mapping $f_1$ to be the identity on $\mathbf{A}$. Since $\mathbf{D}$ is ultrahomogeneous, the isomorphism $f_2$ can be extended to an automorphism. Let us call this automorphism $g$. Now define $g_1$ to be the automorphism $g$ restricted to $B_1$ and $g_2$ to be the identity on $B_2$. This gives us an amalgamation since $g_1 \circ f_1 = g_2 \circ f_2$ on the set $g(B_1) \cup B_2$.




For the proof of part (b) it is enough to show that there exists $\mathbf{D}$ that has the extension property and whose age is $\mathcal{K}$. The homogeneity and uniqueness of $\mathbf{D}$ is ensured by Lemma 1.2.5. $\mathbf{D}$ is again constructed using the union of a chain of finite substructures $\mathbf{D_i}$. For this, a countable number of joint embeddings defined by the **JEP** for $\mathcal{K}$ are used in order to make $\mathrm{Age}(D) = \mathcal{K}$, together with a countable number of amalgamations defined by the **AP** for $\mathcal{K}$ for the construction of each particular $\mathbf{D_i}$. □

This theorem can be used as a tool for finding ultrahomogeneous structures. The identification proceeds as follows. Imagine that you have a countable class $\mathcal{K}$ of structures. Usually it is not an onerous task to show that this class is isomorphism-closed and meets the requirements of **HP**. While **JEP** is usually a special case of **AP**, this is not generally true – e.g. for fields [67].

Since amalgamation class is a central term in the theory of infinite ultrahomogeneous structures, and sometimes it is reasonable to study particular substructures that are "amalgamated", there are several corresponding notions that it is useful to define. Let $\mathbf{A}, \mathbf{B}_1, \mathbf{B}_2$ be relational structures and $f_1, f_2$ embeddings like those in Definition 1.2.3 of property **AP**. An *amalgamation of* $(\mathbf{B}_1, \mathbf{B}_2, \mathbf{A}, f_1, f_2)$ is any triple $(\mathbf{C}, g_1, g_2)$, where $\mathbf{C}$ is a relational structure, $g_1$ an embedding $\mathbf{B}_1 \to \mathbf{C}$ and $g_2$ an embedding $\mathbf{B}_2 \to \mathbf{C}$, such that $g_1 \circ f_1 = g_2 \circ f_2$. Less formally, an amalgamation "glues together" the structures $\mathbf{B}_1$ and $\mathbf{B}_2$ into a single substructure $\mathbf{C}$ such that the copies of $\mathbf{A}$ coincide. A class $\mathcal{K}$ of finite relational structures is called an *amalgamation class* if it has each of the properties **HP**, **JEP** and **AP**.

A structure $\mathbf{A}$ is *generic* for the class $\mathcal{K}$ if it is universal for $\mathcal{K}$ and ultrahomogeneous.

From Theorem 1.2.8 it follows that every countable ultrahomogeneous structure is universal. The converse is not true. Consider for example the countable infinite *random bipartite graph*. This is a countably infinite bipartite graph with both bipartitions infinite and with the property that for any finite subsets $U$ and $V$ of one of the parts one can always find a vertex $z$ in the other part such that every vertex from $U$ and none from $V$ are adjacent to $z$. Using a simple modification of the argument in the proof of Proposition 1.2.6 one can show that this graph is universal for all finite bipartite graphs. However, it is not ultrahomogeneous. Due to the property of graph just mentioned, there has to exist a subgraph similar to that shown in Figure 1.5.

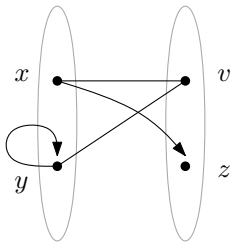



Figure 1.5: A local isomorphism on a subgraph of the random bipartite graph that cannot be extended to an automorphism of the whole graph.

This means that there exist vertices $x, y$ from one part and vertices $v, z$ from the other part where the only adjacent pairs are $\{x, y\}$ and $\{y, v\}$. In this case one can map the subset $\{x, y\}$ onto $\{y, z\}$ using a local isomorphism as indicated in Figure 1.5. This isomorphism cannot be extended to an automorphism, from which we conclude the graph is not ultrahomogeneous.

## 1.3 Combinatorial definitions

This section contains some basic combinatorial definitions needed for later sections providing classifications of graphs and partially ordered sets as well as for the remainder of the work. This section contains virtually no interesting information and may be skipped and referred to only in future whenever notation is not fully understandable.

### 1.3.1 Graph-theoretic definitions

At the beginning of this chapter we defined an undirected graph $G$ as a pair $G = (V, E)$ consisting of vertices $V$ and edges $E$, where edges connect together pairs of vertices. We also defined subgraph and induced subgraph. Similarly homomorphisms and all other mappings are defined accordingly in case of graphs as for relational structures more generally – see also [62].

It is useful to have some notation for several important classes of graphs. The *complete graph* $K_n$ is a graph on $|V| = n$ vertices with all possible edges, i.e. $E = \binom{V}{2}$. For a graph $G = (V, E)$ its *complement* $\overline{G}$ is the graph with vertex

set $V$ and edges $\binom{V}{2} \setminus (\binom{V}{2} \cap E)$. The complement of a complete graph is called an *empty graph*. We define the *degree* $deg_G(v)$ of a vertex $v$ in a graph $G$ as the number of edges of $G$ that contain the vertex $v$. These edges are usually called *incident edges*. Sometimes it is necessary to work with all the vertices connected to a given vertex $v$ by an edge: the *neighborhood* $\Gamma(v)$ is defined as set of all $u \in V(G)$ such that $vu \in E(G)$.

A frequent situation encountered within graphs is the need to connect vertices by a sequence of edges. We can construct a graph $P_n$, called a *path* on vertices $V = \{v_1, v_2, \ldots, v_n\}$, using edges $E = \{v_1v_2, v_2v_3, \ldots, v_{n-1}v_n\}$. We say for vertices $x$ and $y$ that *there exists a path of size* $k$ connecting vertices $x$ and $y$ of $G$ if there is an injective homomorphism $f$ from the path $P_k$ with vertex set $\{v_1, v_2, \ldots, v_k\}$ into $G$ such that $f(v_1) = x$ and $f(v_k) = y$. For a graph $G$ and vertices $v_i, v_j \in V(G)$ there can exist many paths connecting those vertices. If there exists at least one such path, we can define the set of paths joining $v_i$ and $v_j$ that have the shortest length. Paths from this set are called *shortest paths* and their length determine the *distance* $d_{i,j}$ *between vertices* $v_i$ *and* $v_j$. A matrix with all these values defined as $D = (d_{i,j})_{i,j=1}^n$ is called *distance matrix*. Moreover we can define the *eccentricity* $\epsilon_G(v)$ of vertex $v$ in the graph $G$ as the maximum distance of any $u \in V(G)$ from $v$. Taking maximum eccentricity over all vertices defines the *diameter* $diam(G) = \max_{v \in V(G)}\{\epsilon_G(v)\}$.

We call vertices $x, y \in V(G)$ *connected* if there exists $k$ such that there is a path of length $k$ connecting $x$ and $y$. The whole graph $G$ is *connected* if each pair of distinct vertices of $G$ are connected, othewise the graph is *disconnected*. A graph $G$ has a decomposition into connected subgraphs that are maximal with respect to this property [108]. These subgraphs are called the *connected components* of $G$. The term subgraph can be used in two different senses here. First we say that graph $G$ has graph $H$ as a subgraph if there exists homomorphism $f : H \to G$. Alternatively a graph $G$ has graph $H$ as an induced subgraph when there exists an embedding $g : H \to G$. In particular, if there is an embedding $f : K_n \to G$ we call this subgraph $f(K_n)$ a *clique* of $G$ and if there is an embedding $g : \overline{K_n} \to G$ we call $g(\overline{K_n})$ an *independent set* of $G$.

A graph $G = (V, E)$ is *bipartite* if its vertex set $V$ can be decomposed into sets $V_1$ and $V_2$ such that $V = V_1 \cup V_2$ and for each edge $e = \{x, y\}$ we have either $x \in V_1$ and $y \in V_2$, or $x \in V_2$ and $y \in V_1$. We call these parts *partites*. Note that paths are bipartite graphs, although the situation can change when paths are slightly modified. When ends of paths are identified the resulting graphs are called cycles, i.e. the *cycle* $C_n$ of size $n$ on vertex set $V = \{v_1, v_2, \ldots, v_n\}$ has edge set $E = \{\{v_i, v_{i+1}\} : i = 1, 2, \ldots, n-1\} \cup \{v_n, v_1\}$. Only cycles of even size are bipartite – see e.g. [108]. The *girth* of a graph $G$ is the length of the shortest cycle in $G$. The extremal graph with respect to bipartiteness is the *complete bipartite graph* $K_{m,n}$ which is simply a bipartite graph with partites of sizes $m$ and $n$ and with all possible edges between them. Generalizing bipartite graphs, a *multipartite* graph $G = (V, E)$ is a graph whose vertex set $V$ can be decomposed into $m$ sets $V_i$ such that $V = \bigcup_{i=1}^m V_i$ and edges only connect vertices from different partites, i.e. for each edge $e = \{x, y\}$ we have $x \in V_i$ and $y \in V_j$ where $i \neq j$. Again an obvious extremal case is that of the *complete multipartite graph* $K_{n_1,n_2,\ldots,n_m}$ in which all possible edges are present.

Another important class of bipartite graphs are trees. A graph $T$ is called a

*tree* if it is connected and contains no cycles. Vertices of degree 1 are called *leaves*. We call a graph *$H$-free* if it does not contain an induced subgraph of $H$. More generally, for a family of graphs $\mathcal{F}$ we call a graph *$\mathcal{F}$-free* if it does not contain graphs in $\mathcal{F}$ among its induced subgraphs. Denoting the class of all cycles by $\mathcal{C}$, trees are $\mathcal{C}$-free graphs.

We also need the following construction. Given a graph $G$ we define the graph $L(G)$ with vertex set $V(L(G)) = E(G)$ and edge set defined as those pairs of distinct vertices whose corresponding edges in $G$ share a vertex. This graph is called the *line graph* of $G$.

### 1.3.2 Partially ordered sets

A *(non-strict) partially ordered set* (or simply a *poset*) is a pair $(P, \leq)$, where $P$ is a set and $\leq$ is binary relation on $P$ which is reflexive, antisymmetric and transitive. A *a strict partially ordered set* (or simply a *strict poset*) is a pair $(P, <)$, where $P$ is a set and $<$ is irreflexive, antisymmetric and transitive. Sometimes we use the symbols $\preceq$ and $\prec$ to avoid confusion with the usual relation $\leq$ and $<$ on numbers – in this case the corresponding posets are denoted by $(P, \preceq)$ and $(P, \prec)$.

Elements $a, b$ of $P$ are *comparable* if $a \leq b$ or $b \leq a$, otherwise they *incomparable*. A poset where every pair of elements are comparable is called a *linear order* or a *total order* or a *chain*. On the other hand if all elements are incomparable such a poset is called an *antichain*. Given a poset $(P, \leq)$ there may be a *greatest element* $a_g$, which is defined by the property that $a \leq a_g$ for every element $a \in P$. Similarly, there may be a *least element* $a_\ell$, defined by the property that $a_\ell \leq a$ for every element $a \in P$. A *maximal element* $a_M \in P$ is an element for which there is no $a \in P$ such that $a_M < a$. Similarly a *minimal element* $a_m \in P$ is an element for which there is no $a \in P$ such that $a < a_m$.

An important type of partially ordered set for us is a *diamond*, which is a poset formed from a set of pairwise incomparable elements by adding a greatest element and a least element.

### 1.3.3 Digraphs

A *digraph* is an ordered pair $G = (V, E)$, where $V$ is a nonempty set of vertices and $E \subseteq V^2$ is its set of edges. Graphs can be defined as digraphs whose edge relation is symmetric. We can also consider a digraph $G$ to be a graph with *oriented edges*, possibly in both directions, i.e. for an edge from $x \in V(G)$ to $y \in V(G)$ we write $(x, y) \in E(G)$. This type of edge can also be called an *arc*. Similarly to graphs, for a vertex $v$ of a digraph $G$ we define the *out-degree* $deg^+_G(v)$ of vertex $v$ as the number of vertices $u \in V(G)$ such that there is an arc $(v, u) \in E(G)$, and likewise the *in-degree* $deg^-_G(v)$ as the number of $u \in V(G)$ such that $(u, v) \in E(G)$. For a given vertex $v \in V(G)$ the *out-neighborhood* $\Gamma^+(v)$ of vertex $v \in V(G)$ and *in-neighborhood* $\Gamma^-(v)$ of vertex $v \in V(G)$ are defined as set of all $u \in V(G)$ such that $(u, v) \in E(G)$ or $(v, u) \in E(G)$ respectively. Sometimes it is important to stress that a digraph has loops on all vertices. In this case we call such structure a *reflexive digraph* and highlight this property by denoting the such a digraph by $G^\circ$.

Most graph properties can be easily adapted for digraphs. A homomorphism

has to preserve both edges and their orientation. Likewise there are some terms defined for graphs that when carried over to digraphs require preservation of orientation. For example, a *directed cycle* $C_k$ is defined up to isomorphism as the digraph on vertices $\{1, 2, \ldots, k\}$ whose edges are pairs $(i, j)$ such that $(j - i) \equiv 1 \mod k$. A *transitive tournament* $T_k$ is defined up to isomorphism as the digraph on vertices $\{1, 2, \ldots, k\}$ whose arcs are all pairs $(i, j)$ such that $i < j$. A transitive tournament has an *initial vertex* having no incoming edges (except possibly a loop when we consider tournaments with loops allowed) and a *final vertex* having no outgoing edges (again with exception of possible loop when these are allowed).

The following notions are used when working with classes of digraphs as well as graphs. Most of these definitions originate from [62], which is a standard reference for graph homomorphisms. Let $G$ be a digraph and $H$ a subgraph of $G$. A *retraction* of $G$ to $H$ is a homomorphism $r : G \to H$ such that $r(x) = x$ for all $x \in V(H)$. We also say that $G$ *retracts to* $H$. A *core* is a digraph which does not retract to a proper subgraph. If the digraph is such that it has no endomorphism other then identity then it is called a *rigid* digraph.

Considering homomorphisms there is another important notion. Let $G$ and $H$ be digraphs. We write $G \leq H$ when there exists a homomorphism from $G$ to $H$. The relation $\leq$ is not antisymmetric and thus the set of all digraphs equipped with the relation $\leq$ is not a partially ordered set [62]. However, restricting this set to the set of all nonisomorphic cores $\mathcal{C}$ does yield a partially ordered set $(\mathcal{C}, \leq)$.

If for a pair of digraphs $(F, D)$ it is the case that for any $G \in \mathcal{C}$ there is no homomorphism $G \to D$ if and only if there is a homomorphism $F \to G$ then the digraph $D$ is called a *dual of* $F$ in $\mathcal{C}$. A similar definition of duality applies to graphs.

### 1.3.4 Multicolored graphs

Relational structures are quite general and sometimes not very convenient for graph theorists. For this reason several colored versions of graphs are used. Let us first define one type of colored graph in a general way. A structure $\mathbf{G} = (V, S_1, S_2, \ldots, S_\ell, E_1, E_2, \ldots, E_k)$ is called a *multicolored graph* if $V$ is a set of vertices and $\{S_1, S_2, \ldots, S_\ell\}$ is a set of unary relations defined over $V$ and $\{E_1, E_2, \ldots, E_k\}$ is a set of binary relations defined over $V^2$. Of course, the relations $S_i$ correspond to colors of vertices and relations $E_j$ to colors of edges. Homomorphisms between these graphs have to "preserve" colors. Moreover this definition allows several colors to be assigned to a single vertex or single edge. We shall mostly assume both $\ell$ and $k$ to be finite. Note also that this definition is different from edge-colored graphs as usually defined – an *edge-colored* graph is a graph where every edge is assigned a single color.

While analysis of multicolored graphs with an arbitrary distribution of colors among their edges and vertices seems to be hard, it is convenient to use a slightly modified version of multicolored graphs that puts a restriction on what sort of coloring is possible. Let $L$ be a partially ordered set with ordering relation $\preceq$, least element 0 and greatest element 1. An *$L$-colored graph* is an ordered triple $(V, \chi', \chi'')$ such that $V$ is a nonempty set, $\chi' : V \to L$ is an arbitrary function and $\chi'' : V^2 \to L$ is a function satisfying the following:

1. $\chi''(x, x) = 0$; and

2. $\chi''(x,y) = \chi''(y,x)$ whenever $x \neq y$.

The function $\chi'$ colors vertices of $G$, while $\chi''$ colors edges of $G$. The two conditions imposed on $\chi''$ mean that $G$ is without loops and undirected.

A multicolored graph $\mathbf{A} = (V, E, S_1, S_2, \ldots, S_\ell, R_1, R_2, \ldots, R_k)$ can be thought of as an $L$-colored graph $(V, \chi', \chi'')$ where $L = \mathcal{P}(\{1, 2, \ldots, m\})$, where $\mathcal{P}$ denotes the power set operation, with set-inclusion as the ordering relation and $\chi'(x) = \{j : x \in S_j\}$ and $\chi''(x,y) = \{j : \{x,y\} \in R_j\}$.

Intuitively, $\chi'(x) = 0$ means that there are no colors assigned to $x$, and $\chi'(x) = 1$ means that the vertex $x$ is colored by all available colors. Analogously, $\chi''(x,y) = 0$ means that $x$ and $y$ are nonadjacent, while $\chi''(x,y) = 1$ means that the edge $\{x,y\}$ is colored by all the available colors.

To enable studies of homogeneity we must define what counts as a homomorphism for this type of graph. A *homomorphism* between two $L$-colored graphs $(V_1, \chi_1', \chi_1'')$ and $(V_2, \chi_2', \chi_2'')$ is a mapping $f : V_1 \to V_2$ such that

$$\chi_1'(x) \preceq \chi_2'(f(x)) \text{ and } \chi_1''(x,y) \preceq \chi_2''(f(x), f(y)),$$

for all $x$ and $y$ in $V_1$.

We will further distinguish *bicolored graphs* as $L$-colored graphs $G = (V, \chi', \chi'')$ where $\chi'(x) = \emptyset$ (no colors are assigned to vertices) and $\chi'' : V^2 \to L$ is given by $L = (\{\emptyset, \{r\}, \{b\}, \{r,b\}\}, \preceq)$, i.e. edges can be colored by red, blue, or red-blue (both together).

## 1.4 Classification results for homogeneity

Ultrahomogeneous structures attract mathematicians from the perspective of classification. There are several results dealing with different types of structures classifying all ultrahomogeneous structures of a given type. Fraïssé himself was interested in the ultrahomogeneity of $(\mathbb{Q}, <)$. The age of this structure is composed of all finite linear orders. This class is obviously closed under isomorphism and is hereditary. Furthermore the amalgamation property can be easily shown.

This leads us to the classification of countable ultrahomogeneous partial orders provided by Schmerl.

**Theorem 1.4.1** (Schmerl [138])**.** *Any homogeneous partial order is isomorphic to one of the following:*

1. *a countable antichain $(A_n, <)$.*
2. *$(B_n, <) = A_n \times \mathbb{Q}$, where $(a,p) < (b,q)$ if and only if $a = b$ and $p < q$, i.e. an antichain of chains or more precisely a countable union of copies of $(\mathbb{Q}, <)$, where elements in distinct copies are incomparable.*
3. *$(C_n, <)$ where $C_n = B_n$ and $<$ is defined as $(a,p) < (b,q)$ if and only if $p < q$. This can be represented as a union indexed by $(\mathbb{Q}, <)$ of antichains all of the same size.*
4. *$(P, <)$, i.e. the universal ultrahomogeneous partial order for the class of all countable partial orders.*

The first three classes are relatively trivial and the hard part of the proof lies in proving that Age($P$) has the amalgamation property, from which the remaining ones can be simply shown.

Another example of classification is that of all finite ultrahomogeneous undirected graphs without loops, given by Gardiner [50], extending previous results of Sheehan [142].

**Theorem 1.4.2** (Gardiner [50]). *A finite undirected graph without loops is ultrahomogeneous if and only if it is isomorphic to one of the following graphs:*

1. *a disjoint union of complete graphs all of the same size,* $\bigcup_{i=1}^{k} K_n$.
2. *multipartite graphs* $K_{n_1,n_2,\ldots,n_k}$ *with* $n_1 = n_2 = \ldots = n_k$.
3. *the* 5*-cycle* $C_5$.
4. *the line graph* $L(K_{3,3})$.

The proof of this theorem is purely combinatorial using several properties of finite ultrahomogeneous graphs such as bounded diameter or girth, distance transitivity or specific induced graphs on sets of neighbors to iteratively bound the set of possible candidates. Since this class of finite ultrahomogeneous graphs is important in this work we shall refer to it as the *Gardiner's class* or simply as the *Gardiner graphs*, and denote this class by $\mathcal{G}$.

Moving to the infinite, Lachlan and Woodrow extended this classification to all countable ultrahomogeneous undirected graph without loops.

**Theorem 1.4.3** (Lachlan and Woodrow [92]). *Every countable ultrahomogeneous undirected graph without loops is isomorphic to one of the following:*

1. *A graph* $G \in \mathcal{G}$*, i.e. one of the Gardiner graphs.*
2. *A disjoint union of* $m$ *complete graphs all having same size* $n$*, where* $m, n \leq \omega$ *(to ensure distinctness from Gardiner's graphs* $m$ *or* $n$ *should be* $\omega$*).*
3. *The complement of one of the graphs from previous cases.*
4. *The countable random graph* $R$*, i.e. the universal graph for the class of all graphs.*
5. *The universal* $K_n$*-free graph, for a fixed* $n \geq 3$.
6. *The complement to the previous case.*

From the graph-theoretic point of view another interesting classification is that concerning directed graphs. The first results on this topic are represented by the nice work of Lachlan from the mid 1980s [91]. The complete classification of directed graphs was then given by Cherlin in his extensive work from the end of the 1990s [26].

## 1.5 Structures near homogeneity

In the above classifications of ultrahomogeneous structures there are a relatively small number of classes, although the size of particular classes can be infinite. To understand more deeply the principles underlying ulrahomogeneity this work also analyses structures or classes of structures that are in some sense near to ultrahomogeneity. This can be done for finite structures as presented at the beginning of Chapter 6. Nevertheless, while classifications are usually richer for infinite structures an immediate task is to find infinite structures that have properties resembling those of ultrahomogeneous structures and that might be suspected to be easily extended to these ultrahomogeneous counterparts – see the second part of Chapter 6. To achieve such a goal for infinite structures it is more convenient to start with some well-behaved ones that already have some of the required properties. We have already seen that for universal structures there is an embedding into it of any structure younger than itself. We know for example from Proposition 1.2.6 that the Rado graph $R$ is universal and we also know that any ultrahomogeneous structure is universal from Proposition 1.2.7. Still, this requirement has proved to be too general to work with – see Section 6.5. For these reasons, other examples of structures that have many properties resembling those of ultrahomogenous structures are adopted. These are called $\omega$-categorical structures. However to define them several model-theoretic notions have to be introduced.

### 1.5.1 Models and theories

Let $L_1$ and $L_2$ be signatures such that $L_1 \subset L_2$ and suppose that $\mathbf{A}$ is an $L_2$ structure. Then it is possible to turn $\mathbf{A}$ into an $L_1$-structure by forgetting the relations from $L_2 \setminus L_1$. We call such a structure an *$L_1$-reduct of* $\mathbf{A}$, and denote it by $\mathbf{A}|_{L_1}$, or if specification is not necessary simply a *reduct of* $\mathbf{A}$. On the other hand, if $\mathbf{A}|_{L_1}$ is a given structure $\mathbf{B}$ then we call $\mathbf{A}$ an *expansion of* $\mathbf{B}$.

To describe $L$-structures a more formal language needs to be constructed. For the sake of simplicity, we use a less formal definition since we do not directly use the corresponding notions. For more detailed definitions see [67]. Each language is built from a set of variables and several auxiliary symbols, where variables can be considered as temporary labels for elements of the corresponding $L$-structure. Let us for brevity use the following simplified designation. By a string of symbols we mean a progression of symbols for variables, brackets, commas and other auxiliary symbols in the correct order according to the specific use. Since the signature $L$ of a relational structure has no constant or function symbols we can declare every variable to be a *term* (of signature $L$) and nothing else to be a term of $L$. An *atomic formula of $L$* is a string of symbols given by the conditions:

1. for $s$ and $t$ terms of $L$ the string $s = t$ is an atomic formula, and
2. for any $n > 0$ such that there is an $n$-ary relational symbol $R \in L$ and terms $t_1, t_2, \ldots, t_n$ the expression $R(t_1, t_2, \ldots, t_n)$ is an atomic formula of $L$.

Note that we need a new symbol for equality.

A *formula of $L$* is string of symbols, where

1. all atomic formulas are formulas, and

2. if $\varphi$ and $\psi$ are formulas then so are $\neg\varphi$, $\varphi \wedge \psi$, $\varphi \vee \psi$, and

3. if $\varphi$ is formula and $x$ is variable then $\forall x\varphi$ and $\exists x\varphi$ are formulas.

Note again that there are additional symbols for logical operations that need to be included. Moreover, similar to logic, we distinguish free and bounded occurrences of variables in a formula. A *sentence* is then defined as a formula with no free variables. For example, define for a graph $G = (V, E)$ the notion of a *dominating set* $X$ to be a subset of vertices $X \subseteq V$ such that every vertex not in $X$ is adjacent to at least one vertex from $X$. The property of having a dominating set of size $k$ can be defined by the following formula:

$$\exists x_1 \ldots \exists x_k \forall k \bigvee_{1 \leq i \leq k} (E(x_i, y) \vee x_i = y).$$

For defining structures we usually simply use a signature $L$, although to enable a deeper description of structures using formulas it is more suitable to use the notion of a language. A *language* $L_{\infty,\omega}$ can be constructed by including the whole signature $L$ together with some logical symbols, variables and auxiliary punctuation. Terms, atomic formulas and formulas are either defined in the same way as for $L$ or with a little extra technical care – see [67]. The sub-indices in the notation for the language $L_{\infty,\omega}$ indicate that it is permitted to join an arbitrary number of formulas using logical operators, but there can be only finitely many quantifiers in a row. The notion of a language is useful when using formulas and discussing structures in more detail, whereas the signature is used in a simpler way. In any case both notions are used interchangeably in the text according as to which is most appropriate.

Two important types of language are highlighted here. The first is a language $L_{\infty,0}$ with formulas containing no quantifiers, which is called *quantifier-free*. Another type of language is $L_{\omega,\omega}$ where for each formula the logical symbols join only finitely many formulas and the whole formula is also finite. This language is called a *first order language*.

We define languages so as to speak about structures. How are these languages connected with the structures themselves? Any formula is defined for elements of structures through variables used in terms. Substitution of a tuple $\overline{a}$ of elements of $A$ into a term $t$ is denoted by $t[\overline{a}]$, into an atomic formula $\varphi$ by $\varphi[\overline{a}]$ and into a formula $\psi$ by $\psi[\overline{a}]$, as described in [67].

For $L$-structure $\mathbf{A}$ and $\varphi$ a formula defined in the corresponding language, we say that $\varphi$ *is true in* $\mathbf{A}$ or equivalently $\mathbf{A}$ *is model of* $\varphi$ when the statement of $\varphi$ is true in $\mathbf{A}$ and we write

$$\mathbf{A} \models \varphi.$$

This can be generalized to a set of sentences. For a language $L$, a *theory* $T$ is a set of sentences in the language $L$. For a theory $T$ we can have either one or a class of $L$-structures that are all *models of the theory* $T$. For general language $L$ and class $\mathcal{K}$ of $L$-structures, the *theory* $Th(\mathcal{K})$ *of* $\mathcal{K}$ is defined as the class of all sentences $\varphi$ from $L$ such that $\mathbf{A} \models \varphi$ for every structure $\mathbf{A} \in \mathcal{K}$. If all sentences of the theory are first order then we call such a theory a *first order theory*.

### 1.5.2 $\omega$-categorical structures

There can be many models for any given language. An immediate question is whether there are theories that have only one model and if so what can we say about them? The answer to the first question is "yes", which prompts the following definition.

**Definition 1.5.1.** Let $L$ be a language and $T$ be a theory defined in $L$. Then $T$ is called *$\omega$-categorical* if it has exactly one model of cardinality $\omega$ up to isomorphism. Furthermore an $L$-structure is called *$\omega$-categorical* if it has cardinality $\omega$ and it is a model of an $\omega$-categorical theory $T$ defined in $L$.

Note also that finite structures are not $\omega$-categorical. From the perspective of ultrahomogeneity an interesting property of $\omega$-categorical structures is a limitation on the number of orbits of the automorphism group, a result due to Engeler [36], Ryll-Nardzewski [137] and Svenonius [143].

**Theorem 1.5.1** (Engeler, Ryll-Nardzewski and Svenonius)**.** *For a countably infinite first order structure* $\mathbf{A}$ *the following conditions are equivalent:*

1. $\mathbf{A}$ *is $\omega$-categorical.*
2. *The automorphism group of* $\mathbf{A}$ *has only finitely many orbits on $n$-tuples, for every $n \geq 1$.*

This property influences possible ultrahomogeneity since it limits the number of automorphism types for each arity. In addition to the bound on the number of orbits of the automorphism group there is another property making $\omega$-categorical structures similar to ultrahomogeneous ones.

**Theorem 1.5.2** (Appeared in [19])**.** *If* $\mathbf{A}$ *is $\omega$-categorical then it is universal for the class of all countable structures younger than* $\mathbf{A}$*.*

## 1.6 Class of forbidden structures

For a family of finite relational structures $\mathcal{F}$ the class denoted by $\mathrm{Forb}(\mathcal{F})$ consistis of all relational structures $\mathbf{A}$ such that for any $\mathbf{F} \in \mathcal{F}$ there is no embedding $\mathbf{F} \to \mathbf{A}$. Replacing embedding in this definition by homomorphism results in a class that we shall denote by $\mathrm{Forb}_{\mathrm{h}}(\mathcal{F})$. Note also that in general these classes are different, although they can be the same under specific circumstances.

These types of family posses many interesting features. In Theorem 1.4.3 Lachlan and Woodrow showed that the class $\mathrm{Forb}(K_n)$ has an ultrahomogeneous universal graph. On the other hand, Cherlin and Komjáth [28] showed that for any $n \geq 4$ there is no universal countable graph for the class $\mathrm{Forb}(C_n)$. This constitutes a relatively simple yet large family of graphs having no universal graph.

The class $\mathrm{Forb}_{\mathrm{h}}(C_5)$ can however be shown to have a universal graph [87]. There is though no ultrahomogeneous universal graph for this class. Since this class contains graphs with no cycles of length 5 or 3, every pair of distinct non-adjacent vertices can be connected by a path of length either 2 or 3, but not both together. This leads to the existence of two different independent sets of

size 2 in any universal graph, which results in the impossibility of meeting the ultrahomogeneous condition.

There are other examples of families without a universal structure due to Cherlin and Shelah [24]. They make use of a notion called a *near-path* that is a tree created from a path by attaching a single vertex by an edge to any interior vertex of the path. They showed that for any tree $T$ the class Forb($T$) has a universal structure if and only if $T$ is a path or a near-path.

This class is not only used to distinguish the existence of universal structures. It can also generate examples of families that are relatively simply defined but cover a wide class of structures and have interesting universal structures. Namely for graphs there is following result.

**Theorem 1.6.1** (Cherlin, Shelah, Shi [25])**.** *For any finite family $F$ of finite connected graphs there is a universal graph for the class* $\mathrm{Forb}_{\mathrm{h}}(\mathcal{F})$.

## 1.7 Homomorphism homogeneity

Relatively recently Cameron and Nešetřil gave an alternative definition of homogeneity based on various kinds of homomorphism rather than isomorphism [18]. Because there are several types of homogeneity depending on the type of homomorphism used we usually say that a graph $G$ belongs to one of the following classes:

1. **HH** if every homomorphism from a finite induced subgraph of $G$ into $G$ extends to a homomorphism from $G$ to $G$;
2. **MH** if every monomorphism from a finite induced subgraph of $G$ into $G$ extends to a homomorphism from $G$ to $G$;
3. **MM** if every monomorphism from a finite induced subgraph of $G$ into $G$ extends to a monomorphism from $G$ to $G$.

We will also use abbreviated statements like "a graph $G$ is **MH**", "a graph $G$ is $MH$-homogeneous" or "a graph $G$ has the **MH** property", all of which mean that the graph $G$ belongs to class **MH**. Following the notation used in [18, 17], we also use the following general definition.

**Definition 1.7.1.** A structure **A** is said to belong to the class $XY$ if every $x$-morphism from a finite substructure of **A** into **A** extends to a $y$-morphism from **A** to **A** where pairs $(X, x)$ and $(Y, y)$ can be $(I, iso)$, $(M, mono)$ and $(H, homo)$.

Using this notation, **II** stands for the class of ultrahomogeneous structures. Moreover the term homogeneous is used when the classes from Definition 1.7.1 are discussed in general. Most of the claims, proofs and propositions mentioned in this section are taken from [18] which is an excellent source of further information.

### 1.7.1 Homomorphism-homogeneous graphs

There are some trivial observations that can be made immediately. First, both **HH** and **MM** classes are contained in the class **MH**. For undirected graphs, the complete and null graphs are again trivially **HH** and **MM** and thus **MH**. This last statement can be extended to the following proposition.

**Proposition 1.7.1.** *The only finite* **MM** *graphs are complete graphs and null graphs.*

*Proof.* It is trivial that complete and null graphs are **MM**. The other way round needs just a simple argument. Assume that there is a connected graph which is **MM** and not complete. Then there is a pair of distinct non-adjacent vertices $u$ and $v$ and a pair of adjacent vertices $x$ and $y$. A map taking $u$ to $x$ and $v$ to $y$ is a monomorphism and thus can be extended to a global monomorphism by the **MM** property of the graph. However this monomorphism would increase the number of edges. □

Some of the examples of ultrahomogeneous structures are also homogeneous in the sense of homomorphism-homogeneous classes. However, a general statement about the connection between the class of ultrahomogeneous structures and the above-defined classes cannot be easily established. In the countably infinite case the correspondence of homomorphism-homogeneous classes with the ultrahomogeneous is even more interesting – the Rado graph is an example of both an **HH** and an **MM** structure. As we shall see, this example in fact represents a whole family of **HH** structures. Following in the steps of [18], we start with the definition of a property that helps us define these examples.

**Definition 1.7.2.** A graph $G$ has property ** if any finite set of vertices has a common neighbor.

It turns out that this property induces graphs that have $R$ as a spanning subgraph.

**Proposition 1.7.2** (Cameron and Nešetřil [18])**.** *A countable graph contains $R$ as a spanning subgraph if and only if it has the property **.*

*Proof.* To show that the property ** holds for graphs containing $R$ as a spanning subgraph it is enough to observe that it holds in $R$ itself. The finite set from Definition 1.7.2 can be identified with the set $U$ used in the definition of property * in Definition 1.2.2. Moreover property ** cannot be violated by adding edges.

For the converse we have to find an embedding $f$ of $R$ into a given graph $G$ that possesses property **. We iteratively construct mappings $f_k$ using the back-and-forth method. Let $U_k$ denote the set that has been already constructed as the domain of the mapping $f_k$. If $k$ is odd then take vertex $u_i \in V(R)$ where $i = \min\{j : u_j \notin U_k\}$. This vertex is then mapped to a common neighbor of $f_k(U_k)$. Since this set has to have a common neighbor, assuming ** property of $G$, this image exists. When $k$ is even, take a vertex $v' \notin f_k(U_k)$ and choose $u' \in V(R)$ such that for all $u \in U_k$ it holds that $u' \sim u$ if and only if $v' \sim f(u)$. The existence of this vertex is ensured by property * of $R$. □

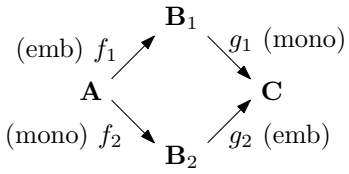
Let property ** hold for a given graph $G$ and let there be a local homomorphism $f : X \to V(G)$. To show that this homomorphism can be extended it is sufficient to show that for any vertex $z \notin X$ there exists a homomorphism $g$ extending $f$. The set $X$ can be divided into neighbors and non-neighbors of $z$. Since we are now using homomorphisms rather than isomorphisms it is enough to require only sufficiently many neighbors of $f(z)$, which is assured by property **. This proves the following proposition.

**Proposition 1.7.3** (Cameron and Nešetřil [18])**.** *Any graph containing $R$ as a spanning subgraph is* **HH** *and* **MM**, *and hence* **MH**.

### 1.7.2 Fraïssé analogue for homomorphism-homogeneity

Proposition 1.7.3 suggests that a direct analogue of Fraïssé's unique limit might be problematic. Nevertheless according to [18] it is at least possible to create an an analogue of Fraïssé's theorem for **MM** relational structures. To achieve that the authors defined the following variant of **AP**.

**Definition 1.7.3.** Let $\mathcal{K}$ be a class of finite relational structures. Then $\mathcal{K}$ has the **mono-amalgamation property** (abbreviated as **MAP**) if for any $\mathbf{A}, \mathbf{B}_1, \mathbf{B}_2 \in \mathcal{K}$ and any maps $f_i : \mathbf{A} \to \mathbf{B_i}$ (for $i = 1, 2$) such that $f_1$ is an embedding and $f_2$ a monomorphism, there exist $\mathbf{C} \in \mathcal{K}$ and monomorphisms $g_i : \mathbf{B_i} \to \mathbf{C}$ for $i = 1, 2$ such that $g_1 \circ f_1 = g_2 \circ f_2$ and $g_2$ is an embedding.




Schematically **MAP** is similar to **AP**, only some embeddings are replaced by monomorphisms. Additional to this property one can define another:

**Definition 1.7.4.** Let $\mathbf{M}$ be relational structure and $\mathcal{K}$ is its age. Then $\mathbf{M}$ has the **mono-extension property** (abbreviated as **MEP**) if for any $\mathbf{B} \in \mathcal{K}$ and substructure $\mathbf{A}$ of $\mathbf{B}$, every monomorphism $\mathbf{A} \to \mathbf{M}$ extends to a monomorphism $\mathbf{B} \to \mathbf{M}$.

This is a variation on the extension property **EP** given by Definition 1.2.4 using monomorphisms instead of isomorphisms. Again it is not very hard to see that an **MM** structure has property **MEP**. Similarly for the converse only a simple variation of Lemma 1.2.5 is required so as to prove the following proposition [18].

**Proposition 1.7.4.** *A countable structure is* **MM** *if and only if it has the mono-extension property.*

Using the previously defined properties and propositions it is possible to prove the following variant of Fraïssé theorem concerning **MM** structures.

**Proposition 1.7.5** (Cameron, Nešetřil [18])**.**

*Let* $\mathbf{D}$ *be a countable* **MM** *relational structure. Then* Age($\mathbf{D}$) *has the mono-amalgamation property* **MAP**.

*Let $\mathcal{K}$ be a non-empty finite or countable set of finite relational structures which has* **HP**, **JEP** *and* **MAP**. *Then there is an* **MM** *relational structure* $\mathbf{D}$ *whose age is $\mathcal{K}$.*

*Proof.* To prove the first part consider the configuration $\mathbf{A}$, $\mathbf{B_1}$, $\mathbf{B_2}$, $f_1$ and $f_2$ defined as in the definition of **MAP**. Similarly to the proof of Theorem 1.2.8 it may be assumed that the structures $\mathbf{B_i}$ are substructures of $\mathbf{D}$ for all $i = 1, 2$ and that the mapping $f_1$ is the identity on $\mathbf{A}$. Again $f_2$ can be extended by a

monomorphism $g$ by the **MM** property of the graph **D**. We use this monomorphism to define an amalgam structure **C** as follows. Let the mapping $g_1$ be the restriction of $g$ to $B_1$ and $g_2$ the identity on $B_2$. Then the domain $C$ would be the union of the domains of structures $g_1(\mathbf{B_1})$ and $\mathbf{B_2} = g_2(\mathbf{B_2})$. This shows that the mono-amalgamation property holds.

To show the existence of an **MM** structure we can construct it iteratively from its age using the given properties, in a similar way to the proof of the original Fraïssé Theorem 1.2.8. Suppose that the structure $\mathbf{D}_i$ has been already constructed. If $i$ is even we choose a structure from $\mathcal{K}$ and use **JEP** to construct a subsequent structure $\mathbf{D}_{i+1}$. Otherwise, if $i$ is odd a pair $(\mathbf{A}, \mathbf{B})$ of structures from $\mathcal{K}$ is chosen such that $\mathbf{A} \subseteq \mathbf{B}$. Since there exists an embedding $f_1 : \mathbf{A} \to \mathbf{B}$, for each monomorphism $f_2 : \mathbf{A} \to \mathbf{D}_i$ we can apply **MAP** as shown in Fig 1.6 and obtain a structure $\mathbf{D}_{i+1}$ such that the monomorphism from **B** to $\mathbf{D}_{i+1}$ extends the monomorphism **A** to $\mathbf{D}_{i+1}$.

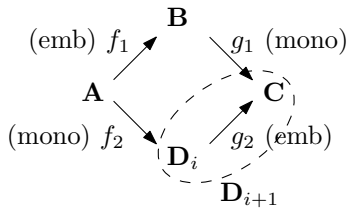



Figure 1.6: Applying **MAP** in the construction of an **MM** structure from its age.

When the steps are arranged in such a way that every structure in $\mathcal{K}$ occurs at an even stage and every pair $(\mathbf{A}, \mathbf{B})$ at infinitely many odd stages, we finally build a countable structure **D**. Moreover every finite substructure of **D** is contained in one of the finite structures $\mathbf{D}_i$ and thus belongs to $\mathcal{K}$. From this one can also see that $\mathcal{K}$ is the age of **D** and has **MAP**, which in turns means that **D** is an **MM**-structure. □

We can see that uniqueness understood as uniqueness up to automorphism is not part of Proposition 1.7.5 and thus we cannot expect to use this result for the construction of a limit as in Fraïssé's original ultrahomogeneous case. It is also evidenced by many examples of **MM** graphs having R as an induced subgraph – the age of every such graph is the class of all finite graphs. However these structures are not completely different – there is at least homomorphism from one to the other.

The question arises whether there is a similar theorem even for homomorphism-homogeneous structures. A class $\mathcal{K}$ of finite relational structures is said to have the **homo-amalgamation property** (abbreviated as **HAP**) if for any $\mathbf{A}, \mathbf{B}_1, \mathbf{B}_2 \in \mathcal{K}$ and any maps $f_i : \mathbf{A} \to \mathbf{B_i}$ (for $i = 1, 2$) such that $f_1$ is an homomorphism and $f_2$ an embedding, there exist $\mathbf{C} \in \mathcal{K}$ and homomorphisms $g_1 : \mathbf{B_1} \to \mathbf{C}$ and embedding $g_2 : \mathbf{B_2} \to \mathbf{C}$ such that $g_1 \circ f_1 = g_2 \circ f_2$. According to recent results in [129] we have the following variant of Fraïssé's theorem for the class **HH** that makes use of the **HAP** property.

**Proposition 1.7.6** (Theorem 4.4 in [129])**.**

*The age of any homomorphism-homogeneous structure has property* **HAP**.

*If a class $\mathcal{K}$ of finite relational structures is isomorphism-closed, has only a countable number of isomorphism classes, and has properties* **HP**, **JEP** *and* **HAP**, *then there is a countable homomorphism-homogeneous structure* **H** *whose age is equal to $\mathcal{K}$.*

## 1.8 Classification results for homomorphism homogeneity

Similarly to the case of ultrahomogeneous structures, homomorphism homogeneity has a relatively extensive classification program. As mentioned earlier, it was already shown in the original paper of Cameron and Nešetřil [18] that **MH** = **HH** for finite graphs and that these classes consist of disjoint unions of complete graphs all having the same size. On the other hand, Proposition 1.7.1 shows that the class of finite graphs with property **MM** consists only of complete and null graphs.

The classes **MH** and **HH** remain coincident for the class of countable undirected graphs, as shown by Rusinov and Schweitzer [136], extending original results of Cameron and Nešetřil [18]. To state this result a term adapted from [136] needs to be defined. For a countable graph $G$ and subset of vertices $T \subseteq V(G)$ we say that $T$ *has a cone*, if there exists a vertex $c \in V(G)$ adjacent to all vertices in $T$.

**Theorem 1.8.1** (Rusinov and Schweitzer [136])**.** *For a countable graph $G$, the following statements are equivalent:*

1. *$G$ is* **MH** .
2. *If $\varphi : H \to H'$ is a monomorphism between finite induced subgraphs $H$ and $H'$ of $G$ which is surjective onto the vertices of $H'$, and $H$ has a cone, then $H'$ also has a cone.*
3. *If $\varphi : H \to H'$ is a homomorphism between finite induced subgraphs $H$ and $H'$ of $G$ which is surjective onto the vertices of $H'$, and $H$ has a cone, then $H'$ also has a cone.*
4. *$G$ is* **HH**

*Proof sketch.* To prove this theorem Rusinov and Schweitzer make use of the just defined notion of a cone. For a monomorphism-homogeneous graph, case 1 can be simply shown. To show the crucial implication between case 3 and 4 they make use of a fact that it is sufficient to show for a local homomorphism $\varphi$ that for any vertex $v \in V(G)$ there is an extension of $\varphi$ by means of introducing $v$ into its domain which is in fact a homomorphism. □

Countably infinite undirected graphs without loops remain interesting since their classification for the classes **MH** and **HH** is incomplete. We have seen in Proposition 1.7.3 that any countable graph containing $R$ as a spanning subgraph is **HH**, **MM** and **MH**. The disjoint union of complete graphs is another

class of homomorphism-homogeneous countably infinite graphs which leads to the question of whether there is another countably infinite **HH** graph which is not isomorphic to either of these cases [18]. This question was answered by Rusinov and Schweitzer several years later [136]. They provided a simple construction for such a graph. Let $n > 1$ and $I_n$ be an independent set of vertices of size $n$. Now for any subset $S_{n-1} \subset I_n$, where $|S_{n-1}| = n - 1$, add an infinite clique and join all its vertices to the vertices in the set $S_{n-1}$. Finally add all edges between any pair of these newly added cliques. This graph is **HH** and does not contain R as a spanning subgraph [136].

Rusinov and Schweitzer in their discussion of the classification of **HH** countably infinite undirected graphs without loops make the observation that some infinite **HH** graphs of this type can be obtained by choosing appropriate finite **HH** graphs with loops and replacing loops with infinite cliques [136]. In this way they were able to give an example of an infinite **HH** graph containing an independent set of size 4 – see Fig. 1.7.

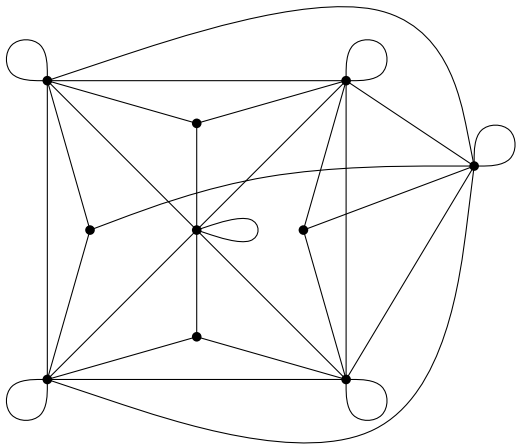


Figure 1.7: An **HH** finite undirected graph with loops that can be used for construction of an **HH** countably infinite undirected graph without loops.

However this observation is not that strong since the following theorem shows that the classification of **HH** graphs with loops allowed seems to be quite difficult.

**Theorem 1.8.2** (Rusinov and Schweitzer [136])**.** *The problem of deciding if a graph with loops allowed is* **HH** *is co-NP complete.*

We have thus seen that for countable graphs classification is still in progress. On the other hand, when we focus on partially ordered sets the situation becomes clearer. Classification of these structures is represented by the two main papers by Mašulovič [109] on the topic, and also two papers by Cameron and Lockett [17]. While Mašulovič's earlier paper is purely devoted to partially ordered sets with nonstrict order, Cameron and Lockett were more interested in classifying all possible classes like **MM**, **MH**, etc. Moreover they were working with partially ordered sets with strict and nonstrict orders. For the sake of brevity, we give only the classification of **HH** countable partially ordered sets.

**Theorem 1.8.3** (Mašulovič [109] and Cameron and Lockett [17])**.** *A partially ordered set* $(A, \leq)$ *is homomorphism-homogeneous if and only if* $(A, \leq)$

1. *is composed of connected components all of which are chains, or*

2. *is a tree, or*

3. *is a dual tree, or*

4. *splits into a tree and a dual tree, or*

5. *is locally bounded and $X_5$-dense.*

For brevity we omit the definition of the new terms here, which can be found in [109].

For directed graphs there are results by Ilić, Mašulović and Rajković determining finite homomorphism-homogeneous tournaments possibly with loops. To state their results several notions have to be defined. Let $Y_n^\circ$ denote the reflexive acyclic tournament on $n$ vertices. Note that such a tournament is unique up to isomorphism. For a given acyclic tournament $Y = (V, E)$ with loops on some vertices, let $P_Y$ denote the set of vertices with loops. For $\emptyset \neq W \subseteq V$, we say that $W$ is *dense in* $Y$ if

1. There exists $x \in P_Y$ such that $(x, w) \in E(Y)$ for all $w \in W$,

2. There exists $x' \in P_Y$ such that $(w, x') \in E(Y)$ for all $w \in W$,

3. For any two vertices $w_1, w_2 \in W$ such that $(w_1, w_2) \in E(Y)$ there exists a vertex $y \in V(Y)$ such that $(w_1, y) \in E(Y)$ and $(y, w_2) \in E(Y)$.

We further call an acyclic tournament $Y$ *dense* if $V(Y)$ is dense in $Y$. Note that $Y_n^\circ$ is a dense acyclic tournament.

**Theorem 1.8.4** ([76])**.** *A tournament $T$ is homomorphism-homogeneous if and only if it belongs to one of the following classes:*

1. *an acyclic tournament with precisely one loopless vertex (this case includes the trivial one-vertex tournament),*

2. *an acyclic tournament with two consecutive loopless vertices and with the additional property that both the initial and the final vertex of the tournament have a loop,*

3. *a dense acyclic tournament,*

4. *the directed cycle $C_3$ or reflexive directed cycle $C_3^\circ$.*

## 1.9 Other variants of homogeneity

There are several other variants of homogeneity that can be considered. For brevity we define most of them just for graphs. These can be simply generalized to relational structures.

One of the most obvious and also oldest variants is to simply add the condition of connectivity to the local isomorphism that is being extended. We call a graph $G$ *connected-homogeneous* if every isomorphism between finite connected induced

subgraphs of $G$ extends to an automorphism of $G$. This notion dates back to 1978 and the work of Gardiner [51] classifying all finite connected-homogeneous graphs. The *Clebsch graph* $\square_5$ is the graph obtained from the 5-dimensional cube $Q_5$ upon identifying antipodal vertices.

**Theorem 1.9.1** (Gardiner [51]). *A finite graph is connected-homogeneous if and only if it is isomorphic to a finite disjoint union of copies of one of the following:*

1. *a complete graph* $K_n (n \geq 1)$;
2. *a multipartite graph* $K_{n,n,\ldots,n} (n \geq 2)$;
3. *a cycle* $C_n (n \geq 5)$;
4. *the line graph of a complete bipartite graph* $L(K_{t,t})(t \geq 3)$;
5. *the Petersen graph;*
6. *the bipartite complement of a perfect matching;*
7. *the Clebsch graph* $\square_5$

This is particularly interesting, since compared to Theorem 1.4.2 classifying finite homogeneous graphs we can see that several homogeneous graphs have been extended to connected-homogeneous families. For example, homogeneous $C_5$ is enlarged to the class of connected-homogeneous $C_n (n \geq 5)$, and homogeneous $L(K_{3,3})$ is enlarged to the class of connected-homogeneous $L(K_{t,t})(t \geq 3)$. Homogeneity of the remaining part of the extended family is violated by the impossibility of extending some local isomorphism on a disconnected set. This could for instance be seen in the case of the Petersen graph. Gardiner also classifies all locally finite connected-homogeneous graphs, which in fact extend the above families of graphs by infinite equivalents of most of the graphs, except for the cases of Petersen and Clebsch graphs, as well as adding two more cases.

This topic is further studied by Lockett in her preliminary results [97] from the perspective of homomorphism-homogeneity. We call a graph $G$ *connected-homomorphism-homogeneous* if every local homomorphism between finite connected induced subgraphs extends to an endomorphism of the graph. For finite graphs the resulting classification is relatively simple. Let us for brevity present only the classification for connected graphs. At first define a graph $\square^2$ as graph created from two $C_4$ by identifying two edges one from each $C_4$.

**Theorem 1.9.2** (Lockett [97]). *Let $G$ be a finite connected graph. Then $G$ is connected-homomorphism-homogeneous if and only if it is one of the following:*

1. *a single vertex;*
2. *a $K_n$-treelike graph $(n \geq 2)$;*
3. *a graph such that all induced cycles are squares, but $\square^2$ has no embedding;*
4. *a bipartite graph such that for each partite $U$ there is a vertex in the other partite adjacent to each vertex in $U$;*
5. *the bipartite complement of a perfect matching.*

There is another extension of classical homogeneity that puts conditions on which isomorphisms need to be extendable. A graph is called *set-homogeneous* if, whenever two finite substructures are isomorphic, there is an automorphism mapping one to another. The main difference is that extension is not required for all isomorphisms. Originally this topic dates back to Enomoto [37], who showed that for finite graphs this notion coincides with classical homogeneity.

## 1.10 Summary

Symmetric structures represented by ultrahomogeneity or its variations constitute a mathematical field interesting from many viewpoints, such as graph theory, combinatorics, group theory and model theory. Moreover there remain many open problems and questions. One of the first type of question to ask is that of classification. Ultrahomogeneous structures based on extending isomorphisms have been widely studied from this perspective, although a complete classification is still not complete [104, 22]. One of the research perspectives as mentioned by Macpherson [104, p. 1605] is to classify homomorphism-homogeneous graphs with edges colored using two or more colors. A basic classification of finite homomorphism-homogeneous bigraphs is given in Chapter 3 and more specifically in Theorem 3.2.5. This warm-up exercise provides a construction called a *pumping argument* that is further used in the more complicated classification of finite homomorphism-homogeneous $L$-colored graphs. Of course, a complete classification of this type would be difficult, so we content ourselves with a classification for the cases where $L$ is a linear partial order or a diamond – see chapter 4.

The last-mentioned result also answers a question about the relationship between the classes **HH** and **MH**. From the very first results of Nešetřil and Cameron [18], followed by Rusinov and Schweitzer [136], these classes were known to coincide for the class of graphs. Their general differences seem to be easily seen, however the question remains: what is a relatively simple structure for which the corresponding classifications result in different classes? An answer is given in Chapter 4, which contains an example of a relatively simply defined type of $L$-colored graph that can be used to distinguish both classes. This is summarized and put into a broader context in Chapter 5, where the general category of classes of Definition 1.7.1 is described and a representative from each class that distinguishes this class from the others is presented.

The search for "relatively simple" structures that distinguish classes **HH** and **MH** motivated another line of research which moreover generalized previous approaches to providing classifications, namely by examining the "distance" of a given structure from ultrahomogeneity as determined by relational complexity – see Chapter 6. This property represents in some sense the simplest structure derived from the original one, having as it does the same automorphism group, and which is ultrahomogeneous. Chapter 6 contains results concerning relational complexity. The first results deal with several classes of finite graphs with small relational complexity like trees followed by a discussion about graphs with high relational complexity. The chapter closes by giving bounds on relational complexity for infinite graphs. For this purpose classes of structures $\mathrm{Forb}_{\mathrm{h}}(\mathcal{F})$ with forbidden homomorphisms are used.

Since one of the motivations for this research stems from the study of symme-

tries of real world networks, especially complex networks like the brain or climate networks, this topic is introduced in Chapter 2 and several results of the author in this field are presented. The main connection lies in the notion of graph limits [100, 118], as explained in Chapter 2.

# 2. Complex network symmetries

> The social psychology of this century reveals a major lesson: often it is not so much the kind of person a man is as the kind of situation in which he finds himself that determines how he will act.
>
> Stanley Milgram

This chapter contains mainly motivational studies performed in the field of complex networks and their connection to the symmetric relational structures studied in the sequel. The reader solely interested in the mathematics of symmetric structures may wish to skip this chapter. Nonetheless, to make the story complete this chapter explains the motivations and connections that has led the author to the deeper study of symmetries of mathematical structures which follows in the remaining chapters.

We begin by describing the field of complex networks itself. Nowadays many real-world systems are often understood as networks, or simply as discrete graphs, of mutually dependent subsystems [121, 13]. The analysis of a specific behavior of these types of systems can be supported by an analysis of this "underlying network", assuming that some of the phenomena are due to a specific connection among its subsystems [121] – for a more detailed discussion see also below. A classical example of a phenomenon that is shared by several networks is the so-called *small-world property*, originating in the famous experiment performed by Stanley Milgram in 1976. Roughly speaking, he showed that for any chosen individual the average distance to any other in terms of friendship, more specifically number of consecutive friendships, is six [114]. The experiment itself was performed in several steps. At first, letters were sent to several individuals at random containing the name of a specified target person. Attached to these letters were instructions that a recipient of the letter should send it either to the target person if he or she knows him or her personally or to the person that he or she knows and considers as the most likely to be connected either with the target person or someone close to him or her by friendship, all these according to the best of the recipent's knowledge. The original algorithm is in fact more complicated – for a more detailed description, see [114]. This notion is sometimes called *six degrees of separation*. This name originated earlier than Milgram's experiement through Hungarian writer Frigyes Karinthy [82]. The surprising aspect of six degrees of separation is that even in systems with a tendency to create groups (like friendship networks) the distance between any pair of participants is relatively small.

Although the assumptions and procedures of Milgram's experiment have been widely discussed and criticized, usually with ajustments suggested in the whole process [139], the small-world phenomenon is a good example of a structural property affecting global behaviour of the whole system. More specifically, it is an example of a simple phenomenon influencing the the whole system from the perspective of connectivity of its subsystems. For a recent review of the small-

world problem see [140].

In more theoretical paper, Duncan J. Watts and Steven Strogatz have shown that similar phenomena can be explained by a simple parametric network model [154]. Their model has two parameters – a probability $p$ and a parameter $k$. This model can be created by a process starting from a graph on $n$ vertices $\{v_1, v_2, \ldots, v_n\}$, each connected to its $k$ nearest neighbors in modular order, i.e. there is an edge $\{v_i, v_j\}$ if and only if $j - i < k \mod n$. This graph is an extremal representation of a phenomenon called *clustering*, which indicates how much a typical vertex neighborhood is dense. However such a graph is not very efficient from the perspective of message delivery through the corresponding network. That is why Watts and Strogatz in [154] propose a way to modify this graph enabling more rapid communication between any pair of vertices by a series of edge-rewirings applied independently with probability $p$.

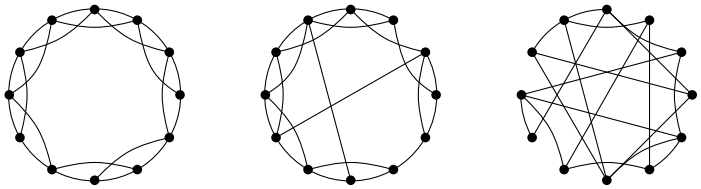


Figure 2.1: Watts and Strogatz parametric model.

Surprisingly, even for small probabilities the resulting graph becomes quite efficient in message delivery while still keeping its clustered nature. This phenomenon has been shown to hold for many real-world networks. For a general reference see [140], for brain networks see [6], and for climate networks see [145] – for a more detailed description of the two latter particular networks see below. It has also been shown that this characteristic can serve as a means of classification, enabling, for example, a potential clinical use in medicine [7].

This simple example shows that the analysis of specific kinds of systems can be supported by studies of the corresponding networks from a graph-theoretical perspective. Since the characteristic of the system under study can be completely represented by a property of the graph structure, one can represent this kind of system as a set of vertices that are connected by edges of a chosen type. This graph is in this case called a *network*. Most of these real-world networks surprisingly share in common such characteristics as the small-world property and therefore a notion *complex network* has been established for them [121].

Another example of a widely shared property is that of a*power-law* degree distribution. To define this notion several auxiliary definitions are needed. The *degree distribution* $P(k)$ is defined as the probability that a node chosen uniformly at random has degree $k$, or equivalently as the fraction of vertices having degree $k$. Let us first consider the finite random graph $G(n, M)$ defined by Erdős and Rényi [38, 39]. This graph $G(n, M)$ has $n$ vertices $v_1, v_2, \ldots, v_n$ and $M$ edges distributed so that the resulting graph is chosen uniformly at random from the family of all graphs which have $n$ nodes and $M$ edges. Alternatively, the Erdös–Rényi random graph can be represented using another closely related model denoted by $G(n, p)$. This model is defined for $n$ vertices $v_1, v_2, \ldots, v_n$, where an edge joining any pair of vertices is included in the graph with probability $p$. Using the second

model the probability that a vertex $v \in V(G(n,p))$ has degree $k$ can be expressed as:

$$P[\deg(v) = k] = \binom{n-1}{k} p^k (1-p)^{n-1-k} \tag{2.1}$$

The average number of edges $z$ connected to each vertex is given by $z = p(n-1)$. The previous equation can then be shown to converge, for large $n$, to

$$\binom{n-1}{k} p^k (1-p)^{n-1-k} \to \frac{z^k e^{-z}}{k!} \tag{2.2}$$

and thus the degree distribution for the random graph $G(n,p)$ is Poisson [123].

Contrary to the previous observation, real-world networks usually exhibit a power law shaped distribution, given by

$$P[\deg(v) = k] = Ak^{-\gamma} \tag{2.3}$$

where the exponent lies in the interval $2 < \gamma < 3$ [3].

Both properties of small-worldness and scale-freeness have become classical characteristics studied for complex networks [121]. For the analysis itself several other graph-theoretic characteristics are also used [13].

Modelling complex networks is a main source of motivation connecting the areas of complex networks and symmetric structures. The key aspect of this motivation is the notion of graph limit [100, 118].

Graph limits as defined by Lovász and Szegedy [101] have been used to analyse a small-world network of coupled Kuramoto oscillators [112]. A different definition of graph limit due to Benjamini and Schramm [8] has been used in the study of the most prominent network-growing model used to represent dynamic complex networks due to Barabási and Albert [3]. Recently Nešetřil and Ossona de Mendez [118] introduced a generalized approach to structural limits based on model theory that enables the definition of limits even for sparser graph sequences than those for which limits had been previously defined. Moreover this type of limit enables the analysis of families of structures for which the countable limit is ultrahomogenous. They showed that if a sequence of graphs converges to an ultrahomogeneous limit than it is convegent under most of the previous definitions of graph limit. This might motivate searching for specific families of complex networks whose limit is ultrahomogeneous. One of the possible properties of the corresponding theory might then be $\omega$-categoricity. These ideas are described in more detail in Section 2.6.

This chapter contains several sections dealing with specific areas of complex networks. In Section 2.1 we describe various types of complex network that are studied and follow this in Section 2.2 with an overview of widely used characteristics and their mathematical definition. Then two specific complex systems are described with an emphasis on results that have been obtained for them – namely these systems are the human brain in Section 2.3 and earth climate in Section 2.4. Both these sections contain a detailed description of the results of the author in these respective fields. Then another result of the author from the study of the small-world phenomeon in a general setting is presented in Section 2.5. Finally, Section 2.6 contains a summary of results and motivation for the joint study of complex networks and symmetric structures.

## 2.1 Network construction

Essential to complex network analysis is the process of constructing a network. Some of the complex systems modeled by networks have their structure defined quite naturally from an underlying physical structure, such as transportation networks like air transportation networks [55] or urban street networks [134, 21]. Computer networks also play an important role in complex network studies. In particular, for the internet several properties similar to complex networks such as its scale-free degree distribution have been established [43]. Analysis of the internet network has been used to assess resilience of the internet to random breakdowns [30]. A more abstract version of this network is that based on the World Wide Web [1, 10] where nodes are HTML pages and edges are determined by references from one page to another. This example is also important because it is generally a directed graph with loops allowed, although loops are sometimes omitted for lack of contribution to connective information. Moreover the World Wide Web should be considered a dynamically changing network, or more accurately a temporal network [70], where a time interval is given and the presence of an edge is given by a collection of time subintervals.

There is another type of system where connections are given by virtual relationships between pairs of vertices from the network represented for example by the frequency of interactions or the amount of shared entities under study. An example of these systems is provided by social networks [153] where connections are characterized by friendships or amount of shared interests – a simple example of a social system is represented by Milgram's small-world experiment. Another type of social system is that of the scientific collaboration network [119], where vertices are scientists and edges are constructed from coauthorships of scientific papers. Here the famous *Erdős number* has been defined as a measure how far an author is from Paul Erdős in sense of co-authorship – see for example [119]. Its definition is simple. Paul Erdős is the only person having Erdős number 0. Everyone who co-authored a paper with him has Edös number 1. Those who have co-authored with someone having Erdős number $i$, but who have not co-authored with any person having a lower value Erdős number, have Erdős number $i + 1$. This may seem more akin to a children's toy, but this or similarly defined characteristics can be a valuable tool for recognizing collaborations in specific fields [119]. To analyse such phenomena another type of network can be used, which has scientific papers as its vertices and (directed) edges where a citation exists in one paper of another – see for example [132].

Social networks and coathorships are today easily accessed via electronic applications like Facebook [95] and Twitter [77] or scientific applications like Web of Science, Scopus, PubMed or Google Scholar [42], and thus construction and interpretation are relatively direct. This field produces or extends many characteristics and phenomena that have been used in other fields – see for example that of centralities [46] or community structures [53]. Other systems consisting of several actors having easily measurable pairwise interactions between themselves are networks based on economic relationships – for example, the world trade network [141], where vertices are countries and edges are given by the existence or amount of imports and exports.

Another example of a complex system for which the physical nature of ele-

ments and their interactions produces the corresponding vertices and their adjacencies is that of chemical reactions/interactions networks, for example the protein-protein network [79] expressing mutual interactions between proteins, or the metabolic network [4] representing the collection of metabolic pathways where enzymes also play a role. This type of network is also interesting since in one of its representations there can be three types of nodes – reactions, reactants and enzymes, and also two types of edge – one for the mass flow of reactants to reactions, and the other for the catalytic role of enzymes in these reactions.

Most of the networks we have mentioned are well-defined by specific phenomena which directly determine the existence of an edge or its weight, although sometimes it can be quite difficult to collect all the necessary information. On the other hand, there are systems where interactions of elements are hidden, due to lack of corresponding data, or even unknown, usually due to the high complexity of interaction. Such networks are constructed using mutual correspondence or influence between particular subsystems (vertices of the network). Edges of these networks are usually defined by a measure of correspondence given by assumptions and conditions connected with system under study. Examples of these networks include brain networks [15], where vertices represent parts of the brain and edges correlation between their activation, climate networks [145], where vertices are geographic locations and edges represents correlation between climatic variables such as temperature or pressure, and financial markets [106, 125], where vertices are stocks traded at specific markets and connections are given by the correlation of their prices or other measures. Such systems should be treated with special care when constructing the complex network model, since many factors can influence the statistical significance of results.

Systems consisting of several subsystems connected by specific kinds of interactions can be sometimes modeled as a collection of oscillators interconnected by synchronization phenomena [2]. However from our point of view the most important part of the system description is always the network.

## 2.2 Characteristics of complex networks

Complex networks are usually constructed because several of their characteristics can be used to identify specific properties, either local or global, of the system they represent. We have already mentioned the small-world phenomenon. This phenomenon becomes more important when considering the dynamics of processes taking place on networks. A good example is given by the dissemination of information. In small-world networks spreading information is much more effective with respect to conveying information from one actor to another [121].

Another widely studied area is that of epidemiological processes taking place through networks [84]. The goal of these studies is usually to predict the spread of diseases by analysing the structure of the underlying networks – for example for sexually transmitted diseases the relevant network is that of sexual partnership [35] or in the case of infectious diseases networks of physical contacts of individuals [113].

To define the small-world phenomenon more precisely we need to introduce several characteristics. For a graph $G = (V, E)$ with $|V| = n$ the *characteristic*

*path length* is computed by the following formula

$$L = \frac{1}{n(n-1)} \sum_{i,j;d_{i,j}<\omega} d_{i,j}, \tag{2.4}$$

where $d_{i,j}$ is the $(i,j)$-entry of the distance matrix of $G$. Note that the graph $G$ may be disconnected and therefore we include only connected pairs of vertices in the summation. This however can cause problems in several applications and therefore sometimes characteristic path length is computed only for a giant component. This characteristic obviously represents reachability of pairs of vertices in the network, with low values indicating high efficiency.

Another term needed is the *clustering coefficient*, defined by

$$C = \frac{1}{n} \sum_i c_i, \tag{2.5}$$

where $c_i$ are *local clustering coefficients* defined for vertex $i$ by

$$c_i = \frac{\sum_{j,\ell} a_{i,j} a_{j,\ell} a_{\ell,i}}{k_i(k_i-1)}, \tag{2.6}$$

where $a_{i,j}$ are elements of adjacency matrix of $G$ and $k_i$ denotes the degree of vertex $i$.

For a vertex $i$ with neighborhood $\Gamma(i)$ and $|\Gamma(i)| = n_i$ and $|E(\Gamma(i))| = m_i$ this local characteristic in fact gives the relative number of edges in the neighborhood of given vertex, i.e.

$$c_i = \frac{|E(\Gamma(i)|}{\binom{n_i}{2}} \tag{2.7}$$

Higher values of the clustering coefficient correspond to a more highly clustered nature of the given network, or less formally, the tendency in the evolving network for groups to be created.

The clustering coefficient also illustrates a general classification principle for characteristics based on their locality. Usually we recognize *local characteristics* as those that are determined by properties of vertices and their neighborhoods while *global characteristics* are determined by features of the whole graph.

In their original paper, Watts and Strogatz [154] suggested the term "small-world" for networks that have similar average path length, but increased clustering coefficient compared to the corresponding random graph, which originally was taken to be the Erdős–Rényi random graph having the same number of edges. More formally, conditions of small-world property can be characterized as

$$\lambda = \frac{L}{L_{rand}} \gtrsim 1, \tag{2.8}$$

$$\gamma = \frac{C}{C_{rand}} \gg 1. \tag{2.9}$$

In real-world applications $L_{rand}$ and $C_{rand}$ are approximated by average estimates $\hat{L}_{rand}$ and $\hat{C}_{rand}$ from several sampled random networks.

This approach, however, has been criticized, since the Erdős–Rényi model is not a precise random model for the corresponding system, because its degree

distribution differs. To overcome this drawback other models, for example such as that of Maslov-Sneppen [107], have been proposed. The Maslov-Sneppen model represents a randomized counterpart for a given network and thus has to be generated with the original network at hand – the Erdős–Rényi model only needs the number of edges and vertices. On the other hand, one can use Maslov and Sneppen model to identify important non-random topological patterns, which are significantly more or less presented in the real network compared to this null-model. The algorithm simply chooses a subgraph $H$ isomorphic to $K_2 + K_2$ at random consisting of two edges $\{x, y\}$ and $\{u, v\}$. Then rewire those edges such that a newly created graph induced on $V(H)$ has edges $\{x, v\}$ and $\{u, y\}$. Maslov and Sneppen suggest repeating this process $4|E(G)|$-times.

The small-world property has been more recently proposed to be summarized in the *small-world index* [74]

$$\sigma = \frac{\gamma}{\lambda} \gg 1. \tag{2.10}$$

Instead of using the path length characteristic, which has problems with disconnected graphs, we can make use of a different measure called *efficiency* [94]. Rather than the arithmetic mean used by the path length characteristic, efficiency is defined using the harmonic mean:

$$E = \frac{1}{n(n-1)} \sum_{i \neq j} \frac{1}{d_{i,j}}. \tag{2.11}$$

Due to the complexity of the underlying system, sometimes it is enough to work with an even simpler characteristic. An example of this is just using degrees of specific vertices or the degree distribution, such as for the definition of scale-freeness. Sometimes a sufficient indicator for specific phenomena can simply be about the connectivity of the vertex, which is here represented by its degree – for an example of using connectivity for clinical purposes in brain studies see [96] and for climate networks see [146].

A more complex property dealing with the general tendency in connectivity as a function of vertex degrees is called *assortativity* [120]. Roughly speaking, this characteristic indicates whether high-degree nodes are likely to be connected with other high-degree nodes, or to low-degree ones. This property is called assortativity, or disassortativity, respectively. Detection of such behavior is usually a nontrivial task, often replaced by calculation of a simple coefficient called the *assortative coefficient* [120], defined by

$$r = \frac{\sum\limits_{(i,j)\in E} k_i k_j - \frac{1}{m}\left[\sum\limits_{(i,j)\in E} \frac{1}{2}(k_i + k_j)\right]^2}{\sum\limits_{(i,j)\in E} (k_i^2 + k_j^2) - \frac{1}{m}\left[\sum\limits_{(i,j)\in E} \frac{1}{2}(k_i + k_j)\right]^2} \tag{2.12}$$

Another class of characteristics dealing with the roles of vertices within a network is called *vertex centrality* [46]. These measures originated in social networks [46], however they were quickly adopted by other fields such as in the analysis of the roles of genes in protein-protein networks [79], detection of energy flows in the climate network [34] or possible patterns of vulnerability of the

brain network in Alzheimer's disease [14]. Generally speaking, centrality indicates important vertices for information exchange and therefore possibly weak places under targeted attacks on networks [69]. One of the most prominent characteristics is *betweenness centrality* [45] defined for vertex $i \in V(G)$ by

$$C_b(i) = \sum_{j,k \in V(G), j \neq i \neq k} \frac{\sigma_{j,k}(i)}{\sigma_{j,k}}, \tag{2.13}$$

where $\sigma_{j,k}$ is the number of shortest paths from vertex $j$ to vertex $k$ and $\sigma_{j,k}(i)$ is the number of shortest paths from vertex $j$ to vertex $k$ going through vertex $i$. Although this characteristic has been widely used it has been criticized for its dependence on shortest paths that sometimes have an unrealistic physical interpretation [122].

## 2.3 The brain network

The brain is a complex system whose structure and function is far from being understood, although it has been known for a relatively long time that the brain consists of neuronal elements interconnected in a complicated network called a *structural network* [144]. Moreover functions of the brain are believed to be better modeled by *functional networks* [47] that consist of several spatially distinct brain regions that are connected according to statistical patterns among them. Both networks can provide interesting insights into brain functionality and both of them can be constructed using various measurement and preprocessing techniques. They are believed to play a role in specific mental processes [48, 110]. This section describes results from [56] dealing with the process of analysis of the brain as a complex network.

### 2.3.1 Defining vertices

In practice the fundamental step is to define network vertices. Usually this definition is tightly connected with the means of measurement used – vertices represent brain regions where activities are measured using for example electroencephalography (EEG), positron emission tomography (PET) or functional magnetic resonance imaging (fMRI). Within these areas activities are measured and the corresponding time series are compared among each other using cross-correlation measures [71].

Although the general theory for complex networks using various methods of measurement, like EEG or fMRI, are similar, there can be some differences in processing. Let us consider functional magnetic resonance imaging as used in [56]. Within this study data from 12 healthy volunteers, 5 males and 7 females in the age range 20 - 31 years, were obtained using a 3T Philips Achieva MRI scanner, operating at ITAB (Chieti, Italy). Each measurement is called a *session*. It has been already mentioned that functional MRI measures activity in specific areas of the brain. The technique called Blood-oxygen-level dependent (BOLD) uses the known coupling between cerebral blood flow or equivalently oxygen level and neural activity in the corresponding area [124]. This is because particular neurons performing tasks are in need of energy which among other things means a higher

amount of oxygen consumption. Oxygen is released to these neurons from the blood at a greater rate compared with inactive neurons. Deoxyhemoglobin, i.e. hemoglobin when carrying no oxygen, is paramagnetic and therefore detectable using an MRI scanner. For more detailed information see [124]. In this way it is possible to observe brain activity as a function of time, however several preprocessing steps should be also applied. Specific preprocessing should be adjusted every time to the task under study, however there are several standard procedures that are applied in most studies. Following [56], we can describe such preprocessing, omitting some measurement-specific algorithms, by:

1. correction for slice-timing differences,
2. correction of head motion across functional images,
3. co-registration of the anatomical image and the mean functional image, and
4. spatial normalization of all images to a standard stereotaxic space, here (Montreal Neurological Institute, MNI) with a voxel size of $3 \times 3 \times 3mm^2$.

For a more detailed description of preprocessing see [49]. For the sake of completeness it should be mentioned that measurements were provided in a so-called *resting state*, which means that subjects perform no task and think about nothing particular, however they do not sleep. Moreover each subject underwent two scanning runs. In each scanning run the initial five dummy volumes allowing the MRI signal to reach a steady state were discarded. The subsequent 300 functional volumes forming a 10-minute data session were then used for the analysis.

Functional MRI usually provides images of the brain at a resolution of 10000 voxels. While a network composed of these voxels can be hard to analyse and results for particular voxels can be hard to interpret, several dimensionality reduction methods are applied. One of the commonly used methods is to map voxels to a standard anatomically designed atlas defining specific areas of interest for the brain – this can be done since we the brain under study will have already been mapped into standard stereotaxic space MNI, see above. One particular automated anatomical labelling (AAL) atlas [149] with 99 parcels, i.e. vertices, was used.

For a more detailed description of preprocessing and the corresponding methods see [56] and references therein.

### 2.3.2 Defining edges

Studies of these kind of systems usually proceed as depicted at Figure 2.2. The main goal of using an association measure is to define the *weighted connectivity matrix*, i.e. a matrix $W = \{w_{i,j}\}_{i=1}^{n}$, where $w_{i,j}$ is either equal to weight of the edge $\{i, j\}$ or equal to 0 when there is no edge. While weighted complex networks in the field of brain have appeared only recently [135] and most studies for brain complex networks are still for unweighted graphs, we apply thresholding to obtain the *connectivity matrix*. However, even determining the weighted connectivity matrix represents a problem in itself. The main issue is hidden in the choice of association measure.

Since recorded fMRI data reflects activities of the corresponding brain regions it is possible to calculate some kind of association measure between the resulting time series. One of the simplest measures is that of *Pearson's correlation coefficient*. For two real random variables $X$ and $Y$ it is defined by

$$\rho(X,Y) = \sqrt{\frac{\mathbb{E}[(X-\mathbb{E}(X))(Y-\mathbb{E}(Y))]}{\mathbb{E}((X-\mathbb{E}(X))^2)\mathbb{E}((Y-\mathbb{E}(Y))^2)}}, \tag{2.14}$$

where the symbol $\mathbb{E}$ is used for expected value. Another option is to use a more complex correspondence measure called *mutual information*, defined for two real discrete random variables $X$ and $Y$ with corresponding sets of values $\Xi$ and $\Upsilon$ by

$$I(X,Y) = \sum_{x\in\Xi}\sum_{y\in\Upsilon} p(x,y)\log\frac{p(x,y)}{p(x)p(y)}, \tag{2.15}$$

where $p(x) = P[X = x], x \in \Xi$ is the probability distribution function of $X$, $p(y) = P[Y = y], y \in \Upsilon$ is the probability distribution function of $Y$ and $p(x,y) = P[(X,Y) = (x,y)], x \in \Xi, y \in \Upsilon$ is the joint probability distribution function for $X$ and $Y$. For continuous variables, mutual information is defined by the corresponding integral. However, in practice mutual information is estimated using discretization of theoretically continuous variables. For a detailed description of the computations involved see [127].

There is a wide range of methods available for the detection of dependence between variables. Pearson's correlation coefficient, a measure particularly sensitive to linear dependence, is widely used. While Pearson's correlation coefficient reliably detects dependence in the case of multivariate Gaussian probability distributions, it may be suboptimal in the case of non-Gaussian dependence patterns. That is why here by linear dependence between two variables is meant a bivariate Gaussian distribution – this peculiar ambiguity is due to naming conventions in the complex systems community.

In the context of curve fitting the term linear dependence is used in a slightly different sense to describe stochastic relationships where the dependent variable is best fitted (e.g. in the least squares sense) by the independent variable using a linear function of the form $Y = aX + \epsilon$, where none of the variables $X, Y$ or noise $\epsilon$ are necessarily Gaussian. This definition includes a broader family of bivariate distributions. However, in the important example of a what is colloquially called "a linear process", that is, stationary multivariate linear autoregressive processes with Gaussian noise, the bivariate dependences among all variables are Gaussian, and therefore "linear" not only in context of curve fitting, but also in our narrower sense.

Usually, interdependence between two variables is calculated because we do not have detailed information about their synchronization – synchronization is itself considered, since variables represent two dynamical (sub)systems with mutual coupling. While synchronization can be quite complex, it is important to consider even nonlinear measures [88].

The question about the suitability of a particular measure for the construction of a complex network is important, since using an inappropriate measure can lead to completely different network characteristics. Let us consider the question more generally and ask

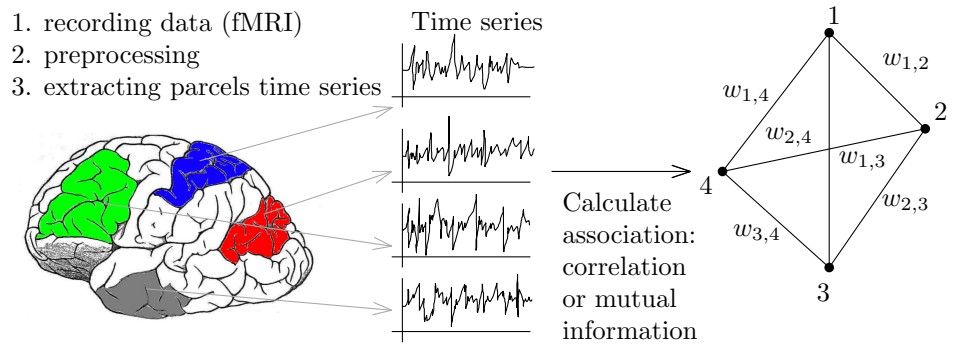



Figure 2.2: Rough scheme for analysis of the brain dynamical system using a complex network approach.

**Problem 2.3.1.** *Is a linear dependence measure represented by Pearson's correlation coefficient when compared to a nonlinear approach represented by mutual information, both computed from fMRI as described above, sufficient to represent the internal dynamics between two regions of the brain in resting state?*

Of course, a complete answer to this question would imply a complete understanding of the underlying system and fMRI measurement and correspondingly the design of a complex analytical solution. However, having in hand measurements of fMRI data, it is possible at least to marshal evidence to guide future research and to construct a method of checking the validity of the chosen method.

Not to speak of the wide application of these chosen measures, these choices are also supported by the property of a bivariate Gaussian distribution that the correlation $\rho(X, Y)$ of the variables $X, Y$ uniquely defines the mutual information between them as [31]

$$I(X, Y) = I_G(\rho(X, Y)) = -\frac{1}{2}\log(1 - \rho^2(X, Y)). \tag{2.16}$$

For a general non-Gaussian bivariate distribution this equation may not hold. Two cases of bivariate non-Gaussianity can be distinguished.

Firstly, "simpler" nonlinearity consists in non-linear rescaling of one or both of the variables. Such a rescaling does not affect the mutual information between the variables, but the correlation may change substantially. Rescaling of this type can be suspected in data e.g. due to non-linear properties of the measurement scale, and may be considered as a bias in the correlation estimation. A remedy commonly adopted is the use of Spearman's rank correlation coefficient. An alternative procedure lies in preprocessing the data by applying a monotonic transformation to each variable separately that would render it Gaussian ("marginal normalization"), and computing the correlation on the transformed data.

A second, more "substantial" type of non-Gaussianity lies in that some bivariate distributions differ from bivariate Gaussian not only in their marginal distributions, but also in the form of the interdependence, which cannot be resolved by univariate rescaling. This substantial non-Gaussianity is the key motivation for the use of nonlinear dependence measures, as the dependence pattern can not be recovered by only considering ranks or other rescaled version of the variables.

Recently, a quantification method for such deviation from Gaussian dependence has been proposed [66], building on the fact that for univariate Gaussian random variables $X, Y$, the correlation gives a lower bound on the mutual information by

$$I(X,Y) \geq I_G(\rho(X,Y)), \tag{2.17}$$

with the minimum obtained for the bivariate Gaussian distribution. In particular, one can define the neglected ("extra-normal" or "non-Gaussian") information

$$I_E(X,Y) = I(X,Y) - I_G(X,Y) \geq 0 \tag{2.18}$$

Considering at least noise, there is some extra information $I_E(X,Y)$ in the data. Our task is to identify whether this amount can have a significant influence on characteristics of the complex network.

### 2.3.3 Linear surrogate datasets

To assess the effect of nonlinearity on the graph-theoretical properties of the resting-state fMRI brain networks we compare the network representations of the data with those of their linear counterparts.

The linear counterpart of a single session dataset is considered as a realization of a linear Gaussian process with the same "linear properties" as the original data. Such realizations, called surrogate datasets, are created using the method of multivariate Fourier transform (FT) surrogates [130, 126]. These are realizations of multivariate linear Gaussian stochastic process which mimic individual spectra of the original time series as well as their cross-spectrum. The multivariate FT surrogates are obtained by computing the Fourier transform (FT) of the series, keeping unchanged the magnitudes of the Fourier coefficients (the spectrum), but adding the same random number to the phases of coefficients of the same frequency bin; the inverse FT into the time domain is then performed. The multivariate FT surrogates preserve the part of synchronization, if present in the original data, which can be explained by a multivariate linear Gaussian stochastic process.

### 2.3.4 Testing of nonlinearity influence

Once the linear surrogate datasets are obtained, we can compute the mutual information for the original possibly nonlinear data and their linear counterparts. The whole process is depicted in Figure 2.3.

The first step is preprocessing and normalization of the original time series from $N_{sess} = 24$ sessions as described above. Then $N_{surr}$ linear surrogates are generated. After this, mutual information for the original data and for all surrogate datasets are computed giving several weighted connectivity matrices. For each matrix several thresholds are applied, resulting in a set of undirected graphs of various densities. These results were used to compute each of the described characteristics from Section 2.2 as a function of density, i.e. $f(\rho)$. To achieve comparable results across the matrices, the function $f = f(\rho)$ is interpolated according to the Voronoi method [85] in which for each value of $\rho$ the value $f(\rho_{neigh})$ is assigned that corresponds to the closest density $\rho_{neigh}$ attained by thresholding.

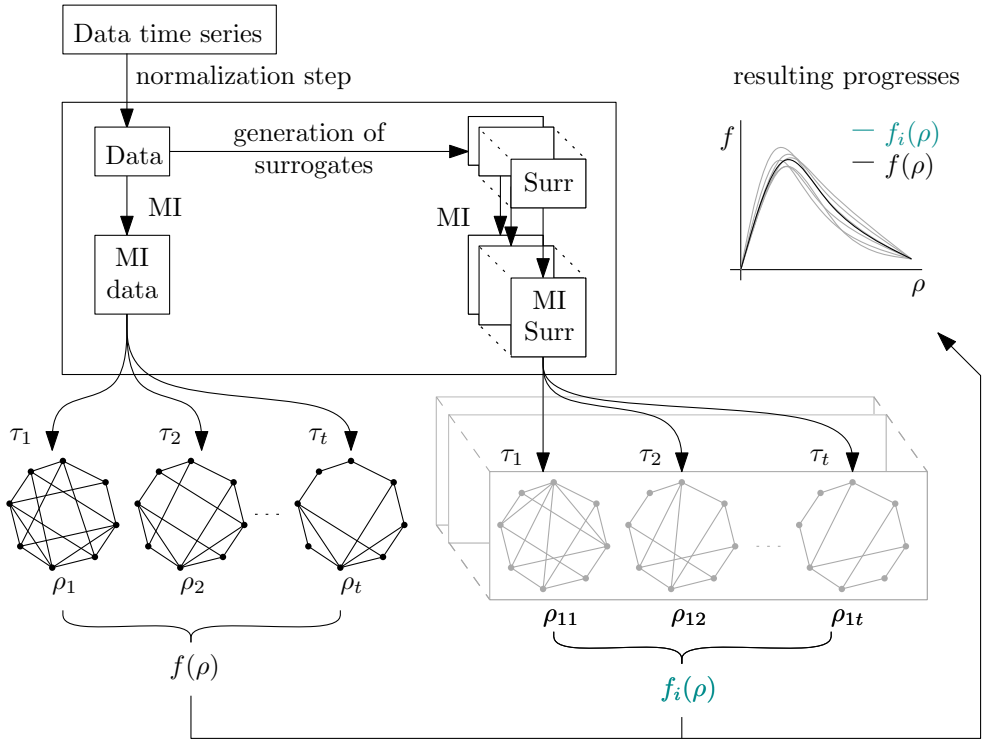
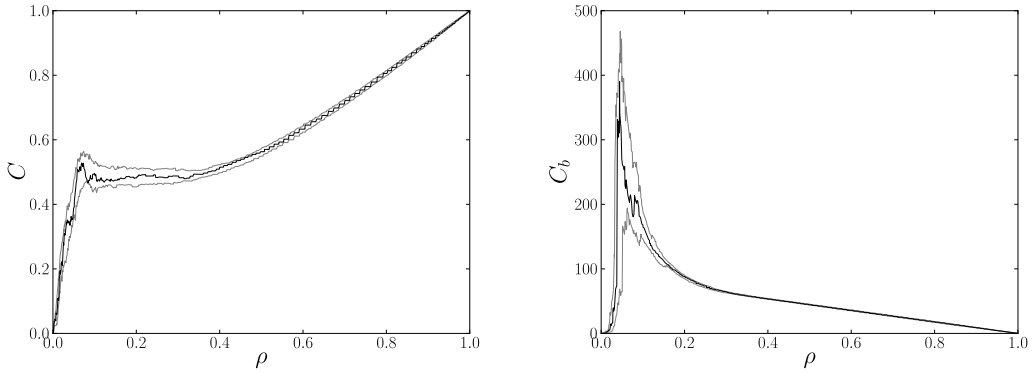



Figure 2.3: Data analysis pipeline: An initial normalization step is performed using the data time series. Then for this data MI is computed for all pairs. In parallel, linear surrogate data are generated, from the original dataset and for each surrogate its MI matrix is computed in a similar way. Each matrix (corresponding to data or surrogate) is then binarized at multiple thresholds providing a collection of graphs. For each such graph a chosen characteristic $f$ is computed and analysed (e.g., plotted as a function of density).

In further description, a data value of the global characteristic for a specific session $s$ and density $\rho$ is denoted by $f^D(s, \rho)$. We drop some of the indices or parameters to simplify notation where confusion is unlikely. We further write $f_j(s, \rho)$ for the values obtained for surrogates, where $j = 1, 2, \ldots, N_{surr}$ are the indices of the surrogates. When local versions of characteristics are used, a parameter $i$ for vertices has to be also used as in $f^D(s, \rho, i)$ and $f_j(s, q, i)$.

For each session, global characteristics are simply computed as functions of density for the data and the corresponding $N_{surr}$ surrogates. The results for surrogates are usually plotted by two gray lines minimal and maximal, forming a "belt" of values typical under linearity. When the black data line falls into this belt, it suggests that the data nonlinearity does not have a substantial effect on the particular graph measure – the deviation from a linear process is within the natural variation due to numerical estimation of the connectivity measure from a finite-size sample. This process is repeated for each session and the goal is to check whether data functions deviate significantly from the "belt" representing the linear surrogate values. For examples of results, see Figure 2.4.

While visual inspection of the belt plot provides an intuitive grasp of the strength of any nonlinearity effect on the studied graph-theoretical measures, a more formal quantitative analysis is beneficial. An intuitive formalization of the




(a) Clustering coefficient $C$ (b) Betweenness centrality $C_b$

Figure 2.4: Two graph characteristics as functions of density $\rho$ for a graph representing original data (black line) and minimal and maximal surrogates lines (gray lines) for a representative session

concept of data 'leaving the belt' of surrogates by the concept of *dominance* is provided below.

Let $s$ and $\rho$ be fixed. For a mesoscopic or local graph characteristic, let us call a vertex $i$ a *maximal dominant vertex* when it holds that $\max_j\{f_j(i)\} < f^D(i)$ and similarly *minimal dominant vertex* if $\min_j\{f_j(i)\} > f^D(i)$. We further define the *maximal dominance indicator function* by

$$f_M^{dom}(s,\rho,i) = \begin{cases} 1, & i \text{ is maximal dominant} \\ 0, & \text{otherwise.} \end{cases} \tag{2.19}$$

The *minimal dominance indicator function* $f_m^{dom}(s,\rho,i)$ is defined in the very same way. Finally, the *dominance function* is defined by $f^{dom}(s,\rho,i) = f_m^{dom}(s,\rho,i) + f_M^{dom}(s,\rho,i)$. The following steps can be performed for any of these functions but we concentrate on the dominance function only. Using simple summation the *graph dominance function* is defined by:

$$f^{G,dom}(s,\rho) = \sum_{i=1}^{n} f^{dom}(s,\rho,i), \tag{2.20}$$

and finally the *overall dominance function* is computed as

$$f^{T,dom}(\rho) = \frac{1}{N_{sess}} \sum_{s=1}^{N_{sess}} f^{G,dom}(s,\rho). \tag{2.21}$$

This function is then further analysed. The same approach is used for global characteristics, with the exception that dominance indicators are not functions of vertices and the summation step (2.20) is therefore omitted. Note that the expected values of the dominances under strict linearity are about 2%, corresponding to the probability of a single value out of a set of 100 values (1 data value and 99 surrogate values) being the highest or the lowest under the simplifying assumption of the existence of a single maximum (or single minimum) value.

To achieve robustness in the statistical comparison of the dominances, these were further averaged across a range of relevant densities. The lower bound is determined so that for all sessions all the resulting graphs are connected, i.e. consist of a single component, so as to guarantee good comparability across graphs. For each session $s$ the minimal density $\rho_{m,s}$ is determined, assuring that the data graph and all the graphs of surrogates are connected.

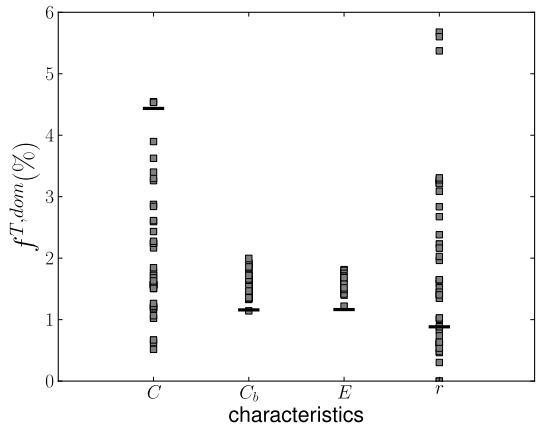



Figure 2.5: Values of global dominances averaged over an interval of interest for shadow datasets (gray dots) and original dataset (black lines) for all mentioned characteristics. Values on the y axis are dominances in percent. Values on the x axis represent the clustering coefficient ($C$), global betweenness centrality ($C_b$), efficiency ($E$) and assortative coefficient ($r$).

To explicitly control for any potential bias in the numerical generation of the surrogate distributions, we repeat for each session the whole procedure for $N_{shadow} = 39$ linear *shadow datasets*. For each session, 39 shadow datasets were created as a multivariate FT surrogate of the marginally normalized original dataset. Thus the shadow dataset preserved only the linear (correlation) structure of the original dataset of the respective session. Subsequently, each shadow dataset has undergone the same procedure as the original data, including the initial normalization, generation of multivariate surrogates, computation of MI and generation of binary graphs using the appropriate thresholds. In this way, we mimic the full procedure of processing the original data while using linear 'shadow' datasets, accounting for any potential slight bias introduced by numerical properties of the algorithm. Then, the dominances for the relevant graph measures computed from the data are compared with the distribution obtained from the linear 'shadow' datasets. In particular, if the data dominance value is higher than at least 38 out of the 39 dominance values obtained from the shadow datasets, this would correspond to a significance level $p < 2/40 = 0.05$ (no correction for multiple comparisons applied).

This procedure results in the plot depicted in Figure 2.5. For all characteristics except efficiency the dominances observed for the original dataset stay within the range of dominances observed for shadow datasets. Dominance of the original dataset for efficiency is outside this interval. However it is a minimum of this set placed not far from the set itself. Such placement would suggest that the dominance effects in shadows were stronger than in the original data which therefore

cannot be caused by nonlinearities (original data more similar to linear surrogates than observed for typical shadow dataset). This more detailed analysis confirmed that nonlinearity does not have a substantial effect on the graph-theoretical properties of the resulting graph, with the data giving similar results as completely linearized shadow datasets.

In the case of local characteristics the situation is slightly more complicated, because dominances are significantly higher than dominances for the shadow datasets. This would suggest some influence of nonlinearities. That is why some additional statistics are applied. Since this section serves more as inspiration for current work, it is suggested that the reader consult [56] for further details. However even these results, similarly to the previously mentioned, suggest that nonlinearity as defined above has no significant impact on complex network characteristics when applied to fMRI brain networks of the described type.

## 2.4 The climate network

Another example of a complex system constituted from several dynamical subsystems that can be analysed using a complex network approach is the climate network [145, 147]. Climatologists work with variables like temperature, atmospheric pressure, wind, precipitation etc. However in contrast to weather, which studies the actual conditions of these variables with possible short-term prediction, climatic studies are concerned more with patterns of variation that emerge over longer time scales and which have long-term effects on weather.

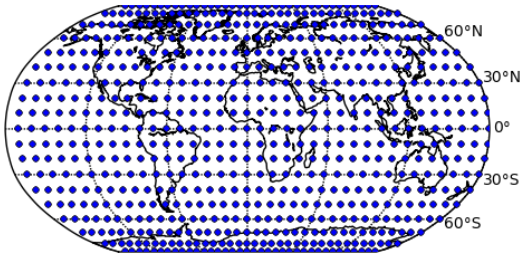
The phenomenon called El Niño-Southern Oscillation (ENSO) [111] serves as a motivational example for using complex networks. This phenomenon involves fluctuation of ocean temperatures in the central and eastern equatorial Pacific. It has a strong influence on specific regional climate patterns in the surrounding areas of the Pacific like North and South America or Australia. There are two patterns with roughly opposite oceanic and atmospheric conditions – the cooler is called *La Niña* and the warmer is called El Niño. Many climatological properties are opposite for these patterns – e.g. during El Niño there are weaker low-level atmospheric winds along the equator, drought in Australia, Indonesia, and neighbouring countries and rains in the central Pacific and the west coast of South America. Moreover in the Atlantic hurricanes tend to be reduced in number and intensity during moderate to strong El Niño events but stronger and more numerous during La Niña events [111]. There can also be periods where neither of these events occur. These events occur irregularly every 2-7 years, although the average is 3-4 years. El Niño typically lasts 9-12 months, and La Niña 1-3 years. This oscillation is usually identified via the *sea surface temperature* (SST), however variations can be detected even in convective rainfall or surface air pressure.

The presence of some of the climatological properties of ENSO, such as changes in predictability [146] or linkage with the Indian summer monsoon [105], has been shown using a complex network approach.

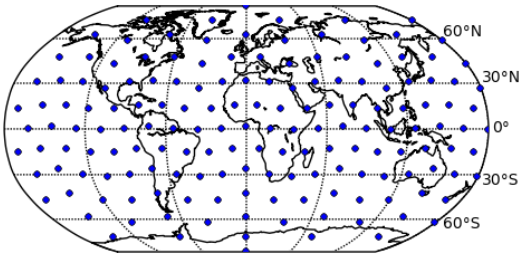
### 2.4.1 Climatic data

Most of the analyses presented in Section 2.4 are based on the NCEP/NCAR reanalysis project [81] which continually updates its dataset using analysis and forecast systems to perform complex data assimilation. These data are usually gridded equidistantly in latitude and longitude 2.5 degrees times 2.5 degrees giving $144 \times 73 = 10512$ time series and possibly in multiple layers representing various pressure levels. Time scales are sub-daily, daily and monthly.

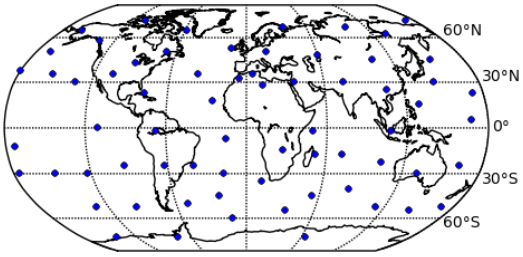
To minimize the bias introduced by periodic changes in the solar input, the mean annual cycle is removed from the data to produce an *anomaly time series*. To enable readability of results several types of dimensionality reductions have been proposed – they are either based on simple sub-sampling of the original latitude longitude grid [148] or conservative remapping schemes for grids in spherical coordinates which assign the same hexagonal area to each grid point [80] or alternatively use a rotated version of *Principal Component Analysis* [65, 5], also known as *Empirical Orthogonal Functions* (EOF) in the climatological community. Various globe samplings resulting in different grids can be seen in Figure 2.6.




(a) Grid defined as sub-sampling of the original NCEP/NCAR grid




(b) Grid defined via remapping to Geodesic Grid where each point corresponds to the same area




(c) Grid defined from data using rotated PCA

Figure 2.6: Various grids defined over the globe for appropriate sub-sampling of the original NCEP/NCAR grid.

The last-mentioned method has the advantage that its use has a long history in climatology [90] and the particular components provided by rotated PCA are usually connected with so-called climatological indices – see for example the overview in [156]. This provides an explanation for most of the vertices and thus subsequent interpretation of complex network results can be simpler. One of the disadvantages is that this dimensionality reduction depends on the data, because PCA uses the analysed data in order to compute the corresponding grid, and therefore comparison of several datasets such as for different variables or

simulation verification can be problematic. On the other hand, if we want to be connected with geographical position, simple sub-sampling or remapping of spherical coordinates seems to be a better choice. The advantage of remapped geographical coordinates lies in the mentioned correspondence of each vertex with the surrounding area always having the same size. Nevertheless simple sub-sampling can more or less fix this handicap by scaling time series by the square root of the cosine of the latitude to account for grid points closer to the poles representing smaller areas and being closer together (thus generating an upward bias of the correlation with respect to grid points farther apart). This method gives grid points mapped to the poles a zero weight, effectively removing them from the analysis.

Using these data the final calculation of association measure as described in Section 2.3.2 can be performed. In this way a network can be constructed and used for climate analysis. Here even a very simple analysis based on identification of areas with high degrees, called *supernodes* [147], provides interesting information about teleconnection patterns – see below.

The term *teleconnection pattern* [151] is used for the correlation of climate characteristics at very distant areas spanning recurring and persistent anomalies over vast geographical regions, sometimes occuring with low frequency. One example of a teleconnection pattern is the *Pacific–North America teleconnection pattern* (PNA) [151], which is a teleconnection pattern with two modes connecting the North Pacific Ocean, represented by the vicinity of Hawaii, and the North American continent, represented as the intermountain region of North America. This pattern was discovered by supernode analysis in the already mentioned study [147].

There is another interesting teleconnection pattern called *North Atlantic Oscillation* (NAO) [5], which binds together the East-Atlantic and West Atlantic again in two phases. Both phases of the NAO are associated with basin-wide changes in the intensity and location of the North Atlantic jet stream. It influences temperature and precipitation patterns from eastern North America to western and central Europe [150]. This pattern cannot be identified from supernode analysis in such direct way as PNA. However when considering the subnetwork consisting of vertices corresponding only to latitudes 30° N the NAO index becomes prominent [147].

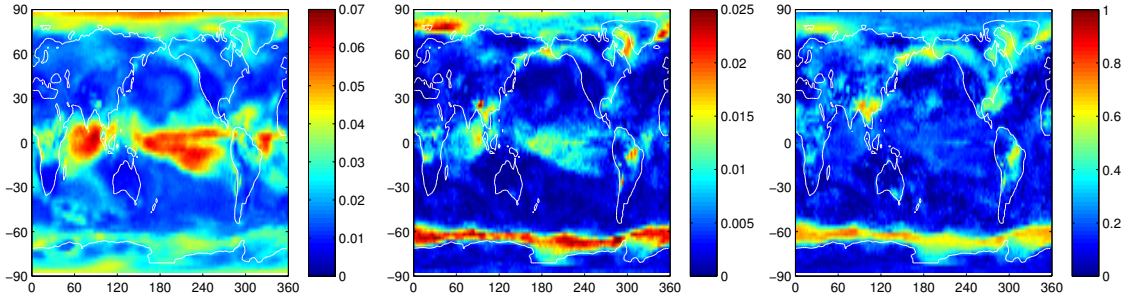
There are other applications in climate networks working with information transfer using the small-world character of the climate network [145] or inferring specific structures using betweenness centrality [34].

### 2.4.2 Connectivity in climate networks

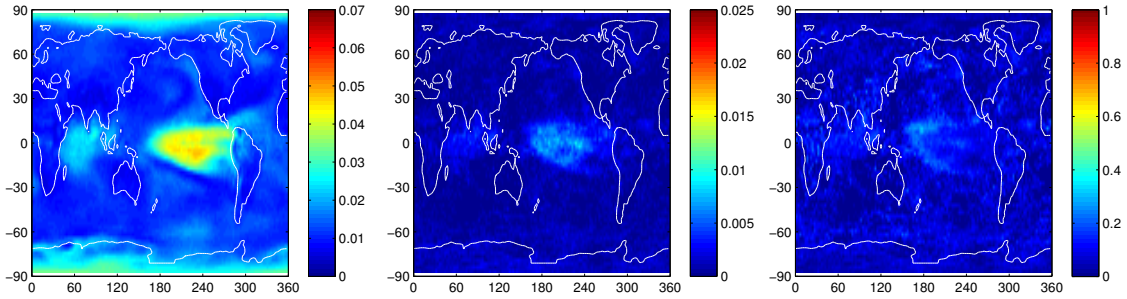
Similarly to brain networks, even the definition of the climate network itself causes many problems. This section presents two problems that have been observed and the corresponding studies designing appropriate corrections.

A study similar in its principles to that presented in Section 2.3.4 has been performed for climate datasets [64], only this time emphasis is put on the influence of nonlinearity on connectivity instead of more complex network measures, since the system under study exhibits the complicated structure of putative nonlinearity. Finally, for most of the alleged nonlinearity an appropriate explanation

and corrective transformations were designed – namely detrending of the original data and consecutive variance normalization [64]. Representative results can be seen in Figure 2.7. It can be observed that most of the nonlinearity in the data can be interpreted using proper preprocessing.



(a) Connectivity for anomalized surface air temperatures (SAT)



(b) Connectity of detrended, variance normalized, anomalized surface air temperatures SAT

Figure 2.7: Sums of weighted degrees representing average contribution of nonlinear dependences in anomalized and detrended variance-normalized data. *Left* average mutual information for each location. *Middle* average nonlinear contribution to mutual information $I_E$. *Right* average nonlinear contribution relative to total mutual information ($I_E/I$)

Apart from nonlinarity there can be another problem represented by a bias in association measures due to the dynamics of the corresponding time series – namely their dynamical memory or higher regularity [128]. The correction is simply achieved, by replacing correlations with Z-scores based on independent Fourier transform surrogate data, which are realizations of processes preserving the original spectra of the studied processes and their entropy rates using Gaussian approximation. For more details see [128].

### 2.4.3 Directed complex networks for climate

Apart from the undirected climate network there are several other extensions or modifications that are valuable in specific kind of analyses. In Section 2.1 we described directed networks used in the analysis of the internet service called the World Wide Web. In the class of complex networks based on the computation of association measures it could also be valuable to consider directed networks so as to make use of causality between events connected to specific vertices.

Such an approach could be particularly relevant due to the potential causal character of processes on the globe given by mass or energy flows. The data

preprocessing is the same as described in Section 2.4.1. However this time we do not compute the symmetric Pearson's correlation coefficient, but apply a different approach that enables us to determine causality between each pair of vertices. We will describe a method applied in the study [65] analysing the reliability of causal measures.

A prominent method for assessing causality is Granger causality analysis, named after Sir Clive Granger, who proposed this approach to time series analysis in a classical paper [54]. However, the basic idea can be traced back to Wiener [155], who proposed that if the prediction of one time series can be improved by incorporating the knowledge of a second time series, then the latter can be said to have a causal influence on the former. This idea was formalized by Granger in the context of linear regression models. In the following, we outline the methods of assessment of Granger causality, following the description given in [33, 52].

Consider two stochastic processes $X_t$ and $Y_t$ and assume they are jointly stationary. Let further the autoregressive representations of each process be:

$$X_t = \sum_{j=1}^{\infty} a_{1j} X_{t-j} + \epsilon_{1t}, \qquad \text{var}(\epsilon_{1t}) = \Sigma_1, \tag{2.22}$$

$$Y_t = \sum_{j=1}^{\infty} d_{1j} Y_{t-j} + \eta_{1t}, \qquad \text{var}(\eta_{1t}) = \Gamma_1, \tag{2.23}$$

and joint autoregressive representation

$$X_t = \sum_{j=1}^{\infty} a_{2j} X_{t-j} + \sum_{j=1}^{\infty} b_{2j} Y_{t-j} + \epsilon_{2t}, \tag{2.24}$$

$$Y_t = \sum_{j=1}^{\infty} c_{2j} X_{t-j} + \sum_{j=1}^{\infty} d_{2j} Y_{t-j} + \eta_{2t}, \tag{2.25}$$

where the covariance matrix of the noise terms is

$$\Sigma = \text{Cov} \begin{pmatrix} \epsilon_{2t} \\ \eta_{2t} \end{pmatrix} = \begin{pmatrix} \Sigma_2 & \Lambda_2 \\ \Lambda_2 & \Gamma_2 \end{pmatrix}. \tag{2.26}$$

The causal influence from $Y$ to $X$ is then quantified based on the decrease in the residual model variance when we include the past of $Y$ in the model of $X$, i.e., when we move from the independent model given by equation (2.22) to the joint model given by equation (2.24):

$$F_{Y \to X} = \ln \frac{\Sigma_1}{\Sigma_2}. \tag{2.27}$$

Similarly, the causal influence from $X$ to $Y$ is defined by

$$F_{X \to Y} = \ln \frac{\Gamma_1}{\Gamma_2}. \tag{2.28}$$

Clearly, the causal influence defined in this way is always nonnegative.

The original introduction of the concept of statistical inference of causality [54] includes a third (potentially highly multivariate) process $Z$, representing all other intervening processes that should be controlled for in assessing the causality between $X$ and $Y$. The bivariate (or "pairwise") implementation of the estimator thus constitutes a computational simplification of the original process, for the sake of numerical stability as well as comparability with the bivariate transfer entropy (conditional mutual information) – for details see [65]. Using this measure it is possible to compute the final directed network as shown in Figure 2.8.

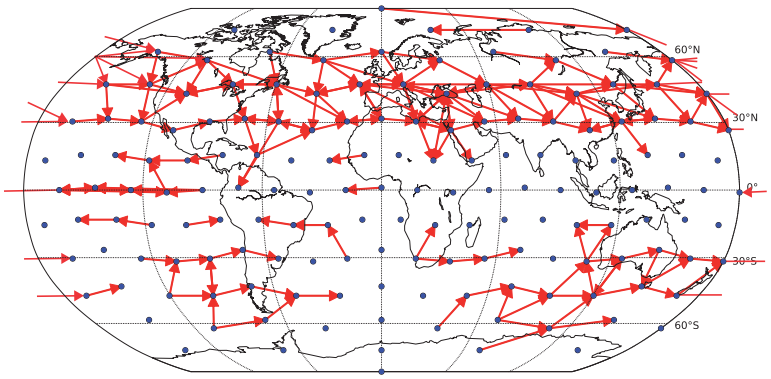



Figure 2.8: Causality network obtained by averaging the results for six decades (total time span 1948–2007) for remapped spherical grid data (162 spatial locations). Only the 200 strongest links are shown. For each decade, the network was estimated by linear Granger causality.

As well as linear Granger causality, nonlinear conditional mutual information has also been used [65]. Results indicate the relative similarity of networks discovered by nonlinear methods with those from linear Granger causality, which is in line with the hypothesis of near-linearity of the investigated climate reanalysis data, in particular the surface air temperature time series.

## 2.5 Small-world in randomly connected systems

The small-world phenomenon as introduced in Section 2.2 is among characteristics widely studied in complex networks. What follows are results from [63] showing that the small-world property can be biased in complex networks constructed from correlation measures.

Some complications with the analysis approach just outlined have recently been suggested in the literature, in particular in the discussion of the biasing effects related to sampling problems such as spatial oversampling [11] and finite size temporal samples [12].

However, we argue that there is a more fundamental problem with the interpretation of the graph-theoretical properties of functional connectivity matrices, i.e., those computed by association measures, especially Pearson's correlation coefficient. In particular, functional connectivity matrices are biased towards a specific structure due to their method of construction, even in the case of perfect sampling. We focus particularly on functional connectivity matrices constructed using linear correlation as a measure of dependence. A specific example of an inbuilt bias of a linear correlation matrix is its 'weak' transitivity property: for any

three random variables $X, Y, Z$ a strong positive correlation between two pairs of them implies a positive correlation within the third pair, including specific 'hard limitations', such as that $\rho_{XY}^2 + \rho_{YZ}^2 > 1$ implies $\rho_{XZ} > 0$, i.e., positivity of the third correlation coefficient (for a proof of a general form of this inequality see [93]). Such implicit dependence between the entries of the correlation matrix have been commonly overlooked in the interpretation of graph analysis of functional connectivity matrices.

We have shown that such dependence has effects on graph structures such as increasing the value of the clustering coefficient and thus affecting small-world indices defined as in equation 2.10.

For simplicity, consider an autoregressive process [102] of order 1 (AR(1)):

$$X_t = c + AX_{t-1} + e_t, \tag{2.29}$$

where $c$ is an $n \times 1$ vector of constants, $A$ is a $n \times n$ matrix whose entries determine the strength of connectivity between particular sub-processes and $e_t$ is an $n \times 1$ vector of error terms. For simplicity we choose $c = \mathbf{0}_{n,1}$ and $e_t \sim \mathcal{N}(0, 1)$ and $A = s(SC + \alpha\mathbb{I})/\lambda_{max}$, with the symmetric binary structural connectivity matrix $SC = SC(n, p)$ generated as a realization of the Erdős–Rényi model $G(n, p)$. In particular, each nondiagonal entry of the $n \times n$ SC matrix is assigned randomly and independently either with value 1 (edge exists, with probability $p$) or value 0 (edge does not exist, with probability $1 - p$); diagonal elements are set to 0. $\mathbb{I}$ here denotes the identity matrix. The parameter $s \in (0, 1)$ modulates the relative strength of the autoregressive and noise terms in (2.29) with the normalization by $\lambda_{max}$ – the largest (in absolute value) eigenvalue of the matrix $SC + \alpha\mathbb{I}$. The autoregressive process is used since many real world systems can be modeled as a system composed of two types of network – structural and functional. For an example of structural and functional networks within brain studies see Section 2.3. Serving as a structural network is that which determines "physical connections" between pairs of subsystems of the global system. On the other hand, the functional network results from the structural by a specific transformation given by the inherent dynamics of the underlying system. This network however determines the overall behavior of the system and it is usually approximated from the data when measuring system outputs. In this case the connectivity matrix of the autoregressive process provides a structural matrix and an AR(1) process provides the required transformation. The resulting time series can be then analysed through the association measure, resulting in a functional matrix. The question is what happens if the structural matrix itself is completely random?

We finish by presenting a motivational example, remaining results being left for inspection in [63]. We generate a finite sample of a stochastic process with length $T = 300$ with parameters $p_{SC} = 0.1, n = 100, s = 0.1, \alpha = 2$, and compute the corresponding functional connectivity matrix $FC$ by binarizing the sample Pearson correlation matrix. The binarization threshold is chosen so that the density $p_{FC}$ of the binarized functional connectivity matrix is equal to the density $p_{SC}$ of the structural connectivity matrix (diagonal elements of the FC matrix are first set to zero). The respective matrices are shown in Figure 2.9.

A visual inspection shows that while the entries of the structural connectivity matrix are independent, the functional connectivity matrix shows a specific

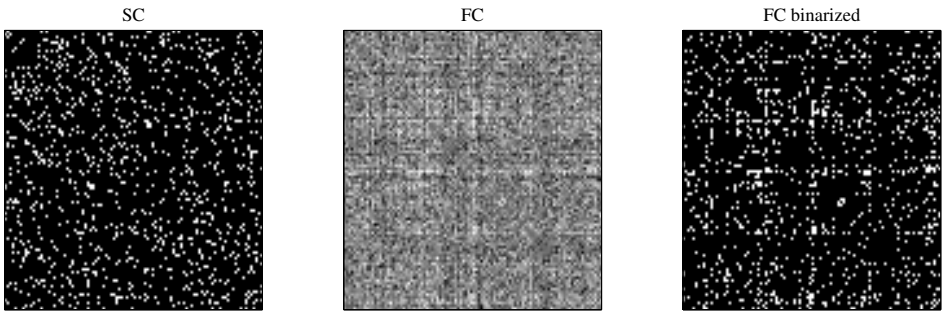



Figure 2.9: Example of a binary functional connectivity matrix (right) generated from a random structural connectivity matrix (left) by thresholding the correlation matrix of the AR-model generated time series (center, light shades of gray indicate higher correlation values). Network with $n = 100$ nodes shown. Note that the functional connectivity matrix shows a specific structure even though the entries of the generating structural connectivity matrix were chosen randomly. See text for further details.

structure. This can be quantified by graph-theoretical measures. In this particular realization, we have $L_S = 2.157, L_F = 2.308, C_S = 0.1081, C_F = 0.2355$. As $SC$ and $FC$ have the same densities, $SC$ is effectively a realization of the Erdős-Rényi model corresponding to the density of $FC$, and we obtain the relative graph measures: $\lambda = 1.07, \gamma = 2.18, \sigma = 2.04$. The values indicate increased clustering and approximately conserved average path length with respect to a corresponding random graph. Together this signifies a small-world-like structure of the functional connectivity matrix, even though the coupling structure of the generating system is completely random.

In view of the results of this simple experiment, it is important to ask what is the significance of the findings of increased clustering and small-world structure in dependence matrices of real-world data. Motivated by this example a series of parametric experiments were performed for various parameters: size of the network $n$, density $p_{SC}$ of the coupling matrix $SC$, balance of the autocorrelation and cross-correlation (by parameter $\alpha$), balance between the autoregressive and noise terms by parameter $s$ and the level of thresholding by changing the required density $p_{FC}$ of the matrix $FC$.

While the strength of the effect is dependent on the length of the sample $T$, the finite size of the sample is not crucial, as the upward bias does not vanish even asymptotically. The covariance matrix $\Sigma$ for the AR(1) process (2.29) is given by the infinite sum $\Sigma = \sum_{i=0}^{\infty} A^i A^{i^T}$ (for a derivation see for example [102]), which due to symmetry of the matrix $A = A^T$ leads to a Neumann series that converges to $\Sigma = (I - A^2)^{-1}$. From this expression the correlation matrix is easily obtained by trivial normalizations using diagonal elements of $\Sigma$. In this way asymptotic behaviour was tested in the parametric study for the mentioned parameters.

In summary, in [63] we showed that the approach of constructing network connectivity graphs from correlations of activity time series of local nodes leads to particular graph structures, characterized by increased clustering compared to common random graph models. This may lead to the attribution of small-

world properties to networks possessing purely random structural connectivity. By extensive mapping of the strength of the effect as a function of the model parameters, we documented that this phenomenon is not a special theoretical case or a negligible effect, but is a rather pronounced and general phenomenon. While the coverage of parameter space was necessarily limited, it is sufficient not only to demonstrate the existence of the effect, but also its potential strength and the existence of complex modulation by system parameters.

For more information see [63] and its supplementary materials.

## 2.6 Concluding remarks and motivation

In the previous sections we have seen various applications of graph theory in the analysis of complex systems via complex network analysis. In particular, we have described the author's work in the analysis of the construction of the complex network for the brain, with the careful handling of nonlinearities in the data outlined in Section 2.3 drawn from the paper [56]. A similarly motivated study of the influence of nonlinearity has been carried out for climate networks, outlined in Section 2.4.2 and further details of which can be found in [64]. In the same section we described the author's other results correcting the bias in the computation of association measures caused by the dynamics of the corresponding time series. This correction is represented by Z-scores based on independent Fourier transform surrogate data and these results can be found in [128]. For climate networks an analysis of the reliability of causal measures was carried out for more complicated directed networks defined via mutual information. This analysis is described in Section 2.4.3 and details can be found in [65]. Finally, an aspect of small-world phenomena was analysed in Section 2.5 for model networks, in which its reliability is under question when computed by correlation measures. Due to the transitive character of this measure, even models generated from completely random subsystems exhibit the small-world characteristic [63].

We move on now to motivating the simultaneous study of complex networks and symmetric structures (especially ultrahomogeneous ones). There is some evidence that complex networks obey some isomorphism type of symmetry – see for example [157, 68, 103, 152]. Although this symmetry is usually defined more locally, it has some influence in improving models of complex networks [157], revealing corresponding identities of a dedicated individual in several different complex systems [158] or analysing specific networks [152].

Another interesting area where symmetries can influence analysis is that of social networks, where the roles of particular participants are studied from the perspective of their similarity [41]. Studies of vertex similarity started with the notion of *structural equivalence* [99], which is simply the property of two vertices having the same neighbourhood, i.e., vertices $u$ and $v$ are structurally equivalent if $\Gamma(u) = \Gamma(v)$. Even more interesting from our point of view is the property called *automorphic equivalence* [41]. Two vertices $u$ and $v$ from a graph $G$ are automorphically equivalent if and only if there exists an automorphism $\pi$ such that $\pi(u) = v$. This notion is interesting, since vertex-transitive graphs represent networks where all vertices have the same role. Moreover, if the definition were extended to groups, or equivalently sets of vertices, ultrahomogeneous graphs are limiting networks in the sense of equivalence between all possible groups.

One of the prominent areas of mathematics that motivates the joint study of symmetric structures and complex networks is that of graph limits [100, 118]. As mentioned by Lovász [100] one motivation for defining graph limits is the understanding and handling of large graph models. To quote one specific example directly he mentions that "Jennifer Chayes, who was studying internet models, asked whether there was a notion of "limit distribution" for sequences of graphs". As well as the internet, Lovász refers to social or biological networks. For these or their models the motivation is to find the corresponding limit. We review here the basis of this story so as to develop this motivation.

Let $G_n$ be a sequence of graphs for which the number of nodes $n$ tends to infinity. Denote by $\hom(F, G)$ the number of homomorphisms from $F$ to $G$. Define the *homomorphism density* as the relative number of homomorphisms

$$t(F, G) = \frac{\hom(F, G)}{|V(G)|^{|V(F)|}}.$$

This value is the probability that a uniformly random mapping $f : |V(F)| \to |V(G)|$ is a homomorphism. We say that a sequence $(G_n)_{n\in\mathbb{N}}$ is *L-convergent*, if the sequence $(t(F, G_n))$ has a limit as $n \to \infty$ for every graph $F$ [101]. This definition is suitable for dense graphs, i.e. those having $\Omega(n^2)$ edges, otherwise it tends to 0. For the very simple example of the random graph with probability $p = 0.5$ it can be shown that corresponding sequence converges with probability 1 [101].

When dealing with graphs whose vertex degrees are at most $D$ another type of convergence called $BS$-convergence is used [8] – in the original paper of Benjamini and Schramm this convergence is called *local weak convergence*. Let $G$ be a graph and $v \in V(G)$. We call a pair $(G, v)$ a *rooted graph*. Any isomorphism $f$ between rooted graphs $(G, v)$ and $(G', v')$ has to satisfy $f(v) = v'$. Denote by $\mathcal{G}_D$ the collection of all isomorphism classes of connected rooted graphs with maximum degree at most $D$. For $r \geq 0$ and $(G, v) \in \mathcal{G}_D$ denote by $B_G(v, r)$ the subgraph of $G$ spanned by the vertices at distance at most $r$ from $v$. Define a metric $\rho$ on $\mathcal{G}_\mathcal{D}$ for rooted graphs as follows. For two rooted graphs $(G, v)$ and $(G', v')$ from $\mathcal{G}_D$ we can find the largest $r$ such that the rooted graph $(B_G(v, r), v)$ is isomorphic to the rooted graph $(B_{G'}(v', r), v')$. Then the metric $\rho$ is defined by $\rho((G, v), (G', v')) = 1/r$. We set $\rho((G, v), (G, v)) = 0$.

A sequence $(G_n)_{n\in\mathbb{N}}$ of finite connected graphs with maximum degree at most $D$ is called *$BS$-convergent* [8] if, for every integer $r$ and every rooted connected graph $(F, o)$ with maximum degree at most $D$ the following limit exists

$$\lim_{n\to\infty} \frac{|\{v : B_{G_n}(v, r) \cong (F, o)\}|}{|G_n|}.$$

We use formulation as given in [118]. Nešetřil and Ossona de Mendez introduce a new notion of limit [118] more generally for any relational structure. Let $\mathbf{A}$ be a finite relational structure with signature $L$. Let $\varphi \in FO(L)$ be a first order formula with free variables $x_1, x_2, \ldots, x_p$. Then let $\Omega_\varphi(\mathbf{A})$ be defined by

$$\Omega_\varphi(\mathbf{A}) = \{(v_1, v_2, \ldots, v_p) \in A^p : \mathbf{A} \models \varphi(v_1, v_2, \ldots, v_p)\}.$$

We define the *Stone pairing* as

$$\langle \varphi, \mathbf{A} \rangle = \frac{|\Omega_\varphi(\mathbf{A})|}{|A|^p}.$$

If the vertices $v_1, v_2, \ldots, v_p$ are chosen uniformly independently at random, this value is the probability that the formula $\varphi$ is satisfied in $\mathbf{A}$. Moreover when $\varphi$ is a sentence the resulting value is either 1 or 0 according to sentence satisfiability.

We call a sequence $(\mathbf{A}_n)_{n\in\mathbb{N}}$ of finite $L$-structures *FO-convergent* if, for every formula $\varphi \in FO(L)$, the sequence $(\langle\varphi, \mathbf{A}\rangle)_{n\in\mathbb{N}}$ is (Cauchy) convergent. *X-convergence* is defined in a similar way when $X$ is fragment of $FO$. For more details see [118].

It turns out that this new limit can simply represent the previously defined ones when adjusted appropriately. Nešetřil and Ossona de Mendez showed that for set $QF$ of quantifier-free formulas $QF$-convergence is equivalent to $L$-convergence.

**Theorem 2.6.1** (Nešetřil and Ossona de Mendez [118])**.** *Let $(G_n)$ be a sequence of finite graphs such that $lim_{n\to\infty}|G_n| = \infty$. Then the following conditions are equivalent:*

1. *the sequence $(G_n)$ is $L$-convergent,*
2. *the sequence $(G_n)$ is $QF$-convergent.*

A similar result can be also given for $BS$-convergence of bounded degree graphs. Let $r \in \mathbb{N}$ and let $\varphi$ be a first order formula with $p$ free variables. Then this formula is called *r-local* if, for every graph $G$ and every $v_1, \ldots, v_p \in G^p$, we have

$$G \models \varphi(v_1, \ldots, v_p) \iff G[N_r(v_1, \ldots, v_p)] \models \varphi(v_1, \ldots, v_p), \tag{2.30}$$

where $G[N_r(v_1, \ldots, v_p)]$ denotes the subgraph of $G$ induced by all the vertices at (graph) distance at most $r$ from one of $v_1, \ldots, v_p$ in $G$. A formula $\varphi$ is local if it is $r$-local for some $r \in \mathbb{N}$. The fragment $FO^{local}$ is the set of all local formulas in $FO$.

**Theorem 2.6.2** (Nešetřil and Ossona de Mendez [118])**.** *Let $(G_n)$ be a sequence of finite graphs with bounded maximum degree such that $lim_{n\to\infty}|G_n| = \infty$. Then the following conditions are equivalent:*

1. *the sequence $(G_n)$ is $BS$-convergent,*
2. *the sequence $(G_n)$ is $FO^{local}$-convergent.*

An interesting situation is when the (countable) limit is unique. This holds for $\omega$-categorical structures. In the case of a sequence of graphs with an ultrahomogeneous limit it has been shown that $FO$-convergence in fact reduces to $QF$-convergence. These and other results presented in [118] when combined with results in [100] suggest that highly symmetric structures may occur naturally as limits of graph sequences.

We already know that real-world networks are modeled by random graph models, especially those exhibiting network growth like for example the Barabási and Albert model [3]. One model and its parametrization may fit complex networks of a specific structure, while another model may be more suitable in different cases. The general problem can be therefore stated roughly as follows.

**Problem 2.6.3.** *Is there a class of complex networks which can be well represented by a network growth model such that this model is FO-convergent and has an $\omega$-categorial structure as its limit?*

A positive solution to Problem 2.6.3 demands that the networks are sparse – see [118]. For the Barabási and Albert model an explicit construction of the $BS$-limit has recently been proposed [9].

# 3. Bicolored graphs

> You take the blue pill – the story ends, you wake up in your bed and believe whatever you want to believe. You take the red pill – you stay in Wonderland, and I show you how deep the rabbit hole goes. Remember, all I'm offering is the truth – nothing more.
>
> Morpheus, The Matrix

This chapter contains a classification of finite homomorphism-homogeneous bicolored graphs. The importance of this chapter lies mainly in the description of a construction called the "pumping argument", used in an extended version later within the classification of finite homomorphism-homogeneous multicolored graphs in Chapter 4. Most of the propositions and theorems presented in this chapter have been published in the paper [61].

In Section 1.3.4 we defined bicolored graphs as a special version of $L$-colored graphs. For the purposes of this chapter it is more convenient to use a definition derived from multicolored graphs. A bicolored graph $G$ will be thus defined as a triple $G = (V, E_1, E_2)$, where $V$ is a set of vertices and $E_1$ and $E_2$ are two binary relations representing colored edges.

The simple classification of homomorphism-homogeneous bicolored graphs contained in this chapter was motivated by efforts to determine simple examples of structures that would distinguish the classes **HH** and **MH**. While for ordinary graphs both classes coincide, we make a start on the classification of a slightly extended version of graphs, namely graphs with colors assigned to edges, and find a representative that separates these classes.

Recall that the finite ultrahomogeneous graphs classified by Theorem 1.4.2 consist solely of disjoint unions of complete graphs all of the same size, complete multipartite graphs with partites of the same size, $C_5$ and $L(K_{3,3})$. Furthermore, recall that the only finite homomorphism-homogeneous graphs [18] are complete graphs and null graphs. We can of course expect that finite homomorphism-homogeneous bicolored graphs will include several colored versions of these graphs, for example by using colors to forbid unsuitable local isomorphisms.

The classification of finite homomorphism-homogeneous bigraphs presented in this chapter can be divided into several steps according to the presence of red-blue edges and weak completeness. All bigraphs in this chapter are finite.

## 3.1 Bigraphs with a red-blue edge

First we consider the case for which there is at least one red-blue edge present.

**Lemma 3.1.1.** *Let $G = (V, E_1, E_2)$ be a homomorphism-homogeneous bigraph with at least one red-blue edge. Then the following holds:*

*(1) Every vertex is incident with a red-blue edge.*

*(2) Let $x, y, z$ be distinct vertices. If $\{x, y\}$ and $\{y, z\}$ are red-blue edges, then $\{x, z\}$ is a red-blue edge.*

*(3) Any two vertices $x$ and $z$ belonging to the same weakly connected component are connected by a red-blue edge.*

*Proof.* To see (1) consider the extension of a partial homomorphism mapping a vertex incident with a red-blue edge to a vertex that is not incident to any red-blue edge.

Part (2) can be derived from the following *pumping argument*. Let $\{x, y\}$ and $\{y, z\}$ be red-blue edges. Assume, for a contradiction, that there is no red-blue edge $\{x, z\}$. Consider a partial homomorphism $f_0$ such that $f_0(x) = x$ and $f_0(z) = y$. Since $G$ is homomorphism-homogeneous $f_0$ can be extended to an endomorphism $f_0'$ over $G$. When extending this local homomorphism an image of vertex $y$ has to be found. We denote this image by $y_1 = f_0'(y)$ and put $y_0 = y$. Vertex $y_1$ has to be connected with vertices $x$ and $y$ by red-blue edges, while $xy_1$ is an image of edge $xy$ under $f_0'$ and $yy_1$ is an image of $zy$ under $f_0'$. For this reason we have $y_1 \neq x$ (assuming no loops) and $y_1 \neq z$ because $\{x, z\}$ is not a red-blue edge. This shows the existence of a new vertex $y_1$ connected by a red-blue edge to both $x$ and $y$.

Another local homomorphism $f_1$ can be defined as an extension of $f_0$ by fixing this new vertex $y_1$. Note that $f_1$ is indeed a homomorphism, since $xz$ can be mapped to $xy$ as before and additionally $zy_1$ can be mapped to $yy_1$. Again, when extending this homomorphism an image of $y$ has to be found and as before the only possible way is to find another vertex $y_2$ connected by red-blue edges to both $x$ and $y$.

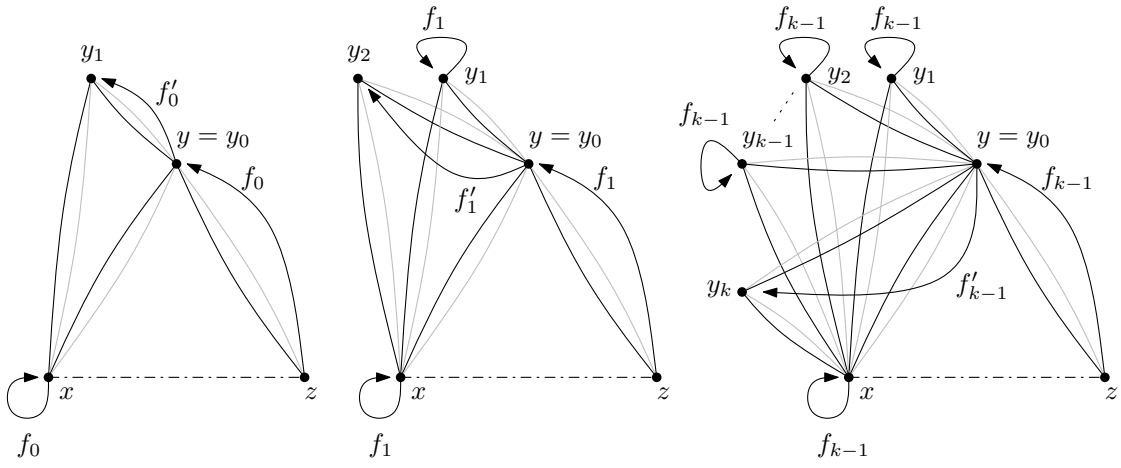



Figure 3.1: The process of extending local isomorphisms by the pumping argument. Blue edges are shown as black and red ones as gray and dash-dotted as a simple chromatic edge or no edge.

Following this procedure (as shown in Figure 3.1) leads to an infinite sequence of vertices $y_0, y_1, y_2, \ldots$. The bigraph $G$ is finite and images cannot be mapped backward, i.e., a mapping $y_j \to y_i$ where $i \leq j$ is impossible. Consequently either

$G$ cannot be homomorphism-homogeneous or there is a red-blue edge $xz$. This leads to a contradiction.

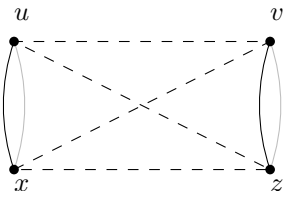



Figure 3.2: Two disjoint vertices $x$ and $y$ in a weakly connected component sharing no neighboring vertex in the bicolored neighborhood. Blue edges are shown as black, red ones as gray, and dashed lines represent a simple chromatic edge.

To show (3) we start with the observation that $x$ and $y$ have to be connected. Otherwise we can consider a homomorphism mapping them into different weak components and extending this local homomorphism is impossible. By (1) there are vertices $u$ and $v$ such that $\{x, u\}$ and $\{y, v\}$ are red-blue edges. If $u = v$, we can apply (2), putting $y = u = v$ to get a red-blue edge $\{x, y\}$.

It remains to consider the case where $u \neq v$. Because any two vertices from a weakly connected component have to be connected, a configuration as depicted in Figure 3.2 must be present. However this configuration is impossible because the local homomorphism $f : \binom{u\ v}{z\ v}$ cannot be extended. This can be shown by using the pumping argument again, where images of vertices $z$ and $x$ are forced to either create an infinite clique or to be mapped backwards to $x$ or $u$, thereby creating a red-blue edge which will then force the red-blue edge $\{x, z\}$. □

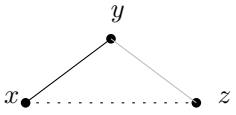
This lemma gives us the following classification of homomorphism-homogeneous bigraphs with at least one red-blue edge:

**Proposition 3.1.2.** *Let $G = (V, E_1, E_2)$ be a finite homomorphism-homogeneous bigraph with at least one red-blue edge. Then every weakly connected component is a complete graph where each edge is red-blue and all these components have the same number of vertices.*

## 3.2 Bigraphs without a red-blue edge

We consider several sub-cases. It is useful to begin with an observation.

**Lemma 3.2.1.** *Let $G = (V, E_1, E_2)$ be a finite homomorphism-homogeneous bigraph that is not weakly complete and has no red-blue edge. Then the following configuration is not an induced subgraph of $G$:*




*where the dotted line represents a non-edge.*

*Proof.* Assume for a contradiction that such a subgraph is induced in the given type of bigraph. The mapping $f : \binom{x\ z}{x\ x}$ is a local homomorphism. Let $f^*$ be its extension. Then connection of $x$ and $y$ by a blue edge and connection of $y$ and $z$ by a red edge implies a connection of $x$ and $f^*(y)$ by a red-blue edge, which yields the desired contradiction. □

The first sub-case we consider is that of bigraphs that are weakly connected but not weakly complete.

**Lemma 3.2.2.** *Let $G = (V, E_1, E_2)$ be a finite weakly connected homomorphism-homogeneous bigraph that is not weakly complete and without red-blue edges. Then the weak diameter of $G$, $d(G)$, is equal to 2.*

*Proof.* Obviously $d(G) \geq 2$ because $d(G) = 1$ holds only for complete bigraphs. For $d \geq 3$ we can find a path of length 3 induced on sequence of vertices $(x, y, z, u)$. The local homomorphism $f : \binom{x\ z}{x\ u}$ cannot be extended, because there is no image for $y$. It follows that $d(G) = 2$ is the only option. □

It follows that there are no further such graphs to consider.

**Proposition 3.2.3.** *There are no finite homomorphism-homogeneous bigraphs without red-blue edge that are weakly connected but not weakly complete.*

*Proof.* Assume for a contradiction that there is a finite bicolored graph $G$ without red-blue edges which is homomorphism-homogeneous. Let $x$ and $z$ be distinct non-adjacent vertices. By Lemma 3.2.2, $d(G) = 2$, so there has to be a third vertex $y$ such that $y$ is adjacent to both $x$ and $z$. Both edges have the same color by Lemma 3.2.1. However for such a situation we can use a slightly modified "pumping argument" like that in Proposition 3.1.1 to reach a contradiction to finiteness.

To do so, assume that both edges $xy$ and $zy$ have the same color, say blue without loss of generality. Similarly to the proof of Proposition 3.1.1, we define a homomorphism $f_0 : \binom{x\ z}{x\ y}$. Since $G$ is by assumption homomorphism-homogeneous, the homomorphism $f_0$ can be extended to an endomorphism $f_0'$. This extension has to map the vertex $y$ to a new vertex $y_1$ which has to be connected by blue edges to $x$ and $y$. We proceed the same way as in the proof of Proposition 3.1.1. Just note that vertices $z$ and $y_1$ cannot be connected by a red edge, since otherwise a homomorphism $g : \binom{x\ z}{x\ x}$ would induce a red-blue edge, which is impossible due to the given assumptions.

Subsequent steps proceed as before by defining the homomorphism $f_1$ the same as $f_0$ for vertices $x$ and $z$ and additionally fixing vertex $y_1$. Due to the character of the connection between $z$ and $y_1$ the mapping $f_1$ is a homomorphism and since $G$ is homomorphism-homogeneous it can be extended to a homomorphism $f_1'$ mapping vertex $y$ into a new vertex $y_2$ connected by blue edges to $x$ and $y$. The remainder of the proof follows the steps made in the proof of Proposition 3.1.1. □

Now consider bigraphs that are weakly connected and weakly complete. In this sub-case all homomorphisms are isomorphisms. Moreover we can view a bigraph $G = (V, E_1, E_2)$ as a graph $H = (V, E_1)$ (considering $E_2$ as non-edges). $G$ is homomorphism-homogeneous if and only if $H$ is ultrahomogeneous. Applying the classification of ultrahomogeneous graphs, we obtain:

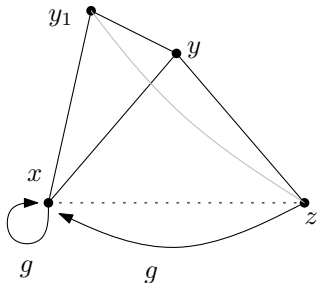



Figure 3.3: Configuration where a monochromatic path of length 2 is extended by a vertex $y_1$, for which an edge $zy_1$ cannot be red, since in such a case extending the homomorphism $g$ to a homomorphism $g'$ would force the existence of a red-blue edge $xg'(y_1)$.

**Proposition 3.2.4.** *Let $G = (V, E_1, E_2)$ be a homomorphism-homogeneous bigraph that is weakly connected, weakly complete and without red-blue edges. Then $G$ is isomorphic to a bigraph that has one of the Gardiner graphs induced on the first color and a complement of the same graph on the second color.*

Finally, we consider bigraphs without red-blue edges that are not weakly connected. By Proposition 3.2.3 it follows that every weakly connected component is weakly complete and all components must be isomorphic. We have already classified such components in Proposition 3.2.4.

Summarizing the above results yields the classification of finite homomorphism-homogeneous bigraphs.

**Theorem 3.2.5.** *A finite bigraph $G = (V, E_1, E_2)$ is homomorphism-homogeneous if and only if it is one of the following:*

1. *a disjoint union of complete graphs all having the same size and composed only from red-blue edges, i.e. $\bigcup_{i=1}^{k} K_n$ for some $k, n \geq 1$,*
2. *a connected bigraph $G_{\mathcal{G}} = (V, E_1, E_2)$ where $(V, E_1)$ is one of the Gardiner graphs and $(V, E_2)$ is its complement,*
3. *a bigraph that has all its connected components isomorphic to bigraphs from the previous case, i.e. $\bigcup_{i=1}^{k} G_{\mathcal{G}}$ for some $k \geq 1$.*

## 3.3 Concluding remarks

All the results in this chapter can be found in [61]. These do not include a resolution of the question of whether the classes **HH** and **MH** coincide for finite bicolored graphs. Considering the proof of Theorem 3.2.5, the only obstacle in proving it also for the class **MH** in the same way as for the class **HH** lies in proving that there cannot be an induced path of length 2 regardless of the color of edges. However getting past such an obstacle would lead to a proof of a variant of Lemma 3.2.1 for **MH** and also the proof of the corresponding variant of Proposition 3.2.3. This should not be a hard task, however it turns out that the coincidence of these classes can be shown in an even more general setting – see Theorem 4.3.5. This latter theorem immediately gives the following proposition.

**Proposition 3.3.1.** *Classes* **MH** *and* **HH** *coincide for finite bicolored graphs.*

Most of the proofs used to obtain this result are not applicable to countably infinite graphs – in fact they make use of finiteness through the pumping argument. This leads us to formulate the following:

**Problem 3.3.2.** *What are the countably infinite homomorphism-homogeneous bicolored graphs?*

Among other reasons this could also be hard because the classification of countable infinite homomorphism-homogeneous graphs is not complete. Moreover, since Proposition 1.7.3 shows that any graph having the Rado graph as a spanning subgraph is **HH**, we can generate an infinite number of bicolored **HH** graphs. We can at least use "sticking", where all non-edges are just converted into the second color. The search for countably infinite homomorphism-homogeneous bicolored graphs should thus be motivated by finding more complex examples.

A solution to the problem of distinguishing the classes **HH** and **MH** that mimics the process for countable undirected graphs by Rusinov and Schweitzer stated in Theorem 1.8.1 seems not to be straightforward due to the several types of edges that can appear within a cone.

Due to the coincidence of the classes **HH** and **MH** for the language of bicolored graphs given by Proposition 3.3.1, the search for a "sufficiently complex" language of structures to distinguish these classes also needs to be moved to the countably infinite case.

**Problem 3.3.3.** *Do the classes* **MH** *and* **HH** *coincide for countable bicolored graphs?*

Another question deals with a specific type of coloring of edges. For any countable bicolored graph $G = (V, E_1, E_2)$ one can find a countable graph $G' = (V, E_1', E_2', E_3)$ where $E_1' = E_1 \setminus E_2$, $E_2' = E_2 \setminus E_1$ and $E_3 = E_1 \cap E_2$, i.e. single colored edges are preserved and those having two colors are colored by a new one. Let us call this process *chromatic unification*. Such a process produces simply edge-colored graphs.

These two versions of colored graphs do not differ at all, or only a little when using isomorphism and speaking about the class **II**, however they can be different when going into the analysis of homomorphisms and the corresponding homogeneity classes. In fact there is a simple example of bicolored graph and its chromatic unification which differs in these properties. It is a specific bicoloring $G$ of $K_4$ and its chromatic unification $G'$ as shown in Figure 3.4.

The illustrated monochromatic unification $G'$ can be easily checked to be homomorphism-homogeneous. However the original bicolored graph $G$ is not homomorphism-homogeneous, as can be shown by taking the local homomorphism $f : \binom{b\ c}{b\ d}$. This shows that allowing multiple colors differs from adding colors, however in this case it is more restrictive in the sense that allowing multiple colors for edges leads to a less symmetric structure in the sense of the class **HH**. This motivates the following problem:

**Problem 3.3.4.** *Is there a countable bicolored graph which is homomorphism-homogeneous, but its chromatic unification is not?*

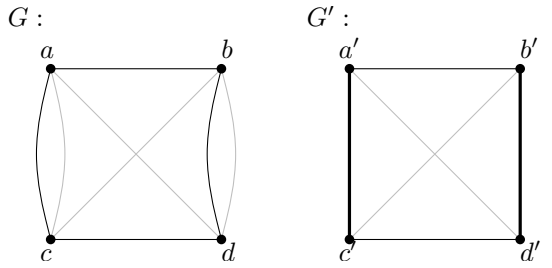



Figure 3.4: Bicolored graph $G$ and its monochromatic unification $G'$. The thinner black edges represent edges with color blue, gray edges with color red and heavier black edges represent those colored green – a new color defined by monochromatic unification.

An obvious generalization of this classification, which leads to other questions, is to add colors to edges or even to vertices. These types of structures are analysed in Chapter 4 which follows.

# 4. $L$-colored graphs

> Mere color, unspoiled by meaning,
> and unallied with definite form,
> can speak to the soul in a
> thousand different ways.
>
> Oscar Wilde

This chapter is motivated by the results of Chapter 3. It seems that introducing only two colors, represented by structures called bicolored graphs, could result in new examples of **HH** structures. Although there are still some open problems for bicolored graphs, see Section 3.3, the natural generalization based on introducing more colors for edges or even for vertices seems to be promising for the identification of new homomorphism-homogeneous structures.

The whole process is motivated by the general though vaguely defined question of extending relational structures that lead to homogeneity – here more precisely to homomorphism-homogeneity.

**Problem 4.0.5.** *How much do simple structures like ordinary graphs have to be extended in order to provide interesting examples of homomorphism-homogeneous structures?*

In the previous chapter we proved that for bicolored graphs such examples consist only of bicolored graphs, where each connected component is a bicolored complete graph, with edges of one color forming a Gardiner graph and edges of the other color forming its complement – see Theorem 3.2.5. In this chapter we propose using more colors for edges and adding the possibility of coloring vertices with further colors. Again it is possible to add a set of colors to each edge or vertex. While such a generalization could create relatively general structures we restrict ourselves just to those cases where all the sets attached to edges constitute a specific partial order under set inclusion.

Another motivation for this chapter is that of the relationship between classes **HH** and **MH** discussed at the end of Section 3. While the language of bicolored graphs seems not to be "strong enough" to distinguish them (the full argument for this is given in this section), a possible extension might be the language of $L$-colored graphs. The motivation is thus to find the simplest example making both classes distinct.

The results of this chapter are contained in [57], with the exception of some of the concluding remarks.

## 4.1 MH-homogeneous $L$-colored graph

In this chapter we propose the study of the notion of $L$-colored graph introduced in Section 1.3.4, together with the corresponding notion of homomorphisms between these graphs. For the purposes of this chapter several other notions have to be introduced.

For $W \subseteq V$, the *substructure of* $(V, \chi', \chi'')$ *induced by* $W$ is $G[W] = (W, \chi'|_W, \chi''|_W)$, where $\chi'|_W$ and $\chi''|_W$ denote the restrictions of $\chi'$ and $\chi''$ to $W$, respectively. For an $L$-colored graph $G = (V, \chi', \chi'')$ and $\alpha \in L$ let $W_\alpha = \{x \in V : \chi'(x) = \alpha\}$ and $G^{(\alpha)} = G[W_\alpha]$.

Note that definitions of homomorphism-homogeneity and monomorphism-homogeneity are the same as defined in Definition 1.7.1, only where the appropriate homomorphism of $L$-colored graphs has to be used instead.

Let $G = (V, \chi', \chi'')$ be an $L$-colored graph, and let $\theta_G \subseteq V^2$ be the reflexive transitive closure of $\theta_G^0 = \{(x, y) \in V^2 : \chi''(x, y) \neq 0\}$. Then $\theta_G$ is an equivalence relation on $V$ whose equivalence classes will be referred to as *connected components* of $G$. An $L$-colored graph $G$ is *connected* if $\theta_G$ has only one equivalence class. Otherwise, it is *disconnected*. We say that $G$ is *complete* if $\chi''(x, y) \neq 0$ for all $x \neq y$.

An $L$-colored graph $G = (V, \chi', \chi'')$ is *vertex-uniform* if there exists an $\alpha \in L$ such that $\chi'(x) = \alpha$ for all vertices $x$, and it is *edge-uniform* if there exists a $\beta \in L \setminus \{0\}$ such that $\chi''(x, y) = \beta$ for all vertices $x$, $y$ such that $x \neq y$. We say that an $L$-colored graph $G = (V, \chi', \chi'')$ is *uniform* if it is both vertex-uniform and edge-uniform. Up to isomorphism, a finite connected uniform $L$-colored graph is uniquely determined by $n = |V|$, the color of vertices $\alpha$ and the color of edges $\beta \succ 0$, and we denote it by $U(n, \alpha, \beta)$.

If there is no danger of confusion, we shall write simply $\chi(x)$ and $\chi(x, y)$ instead of $\chi'(x)$ and $\chi''(x, y)$, respectively. Also, the set of vertices of $G$ will be denoted by $V(G)$.

**Lemma 4.1.1.** *Let $G$ be an* **MH***-homogeneous $L$-colored graph. Assume that there exist three distinct vertices $a_0, a_1, x \in V(G)$ such that:*

*(i)* $\chi(a_0, a_1) \succ 0$ *and* $\chi(x, a_1) \succ 0$,

*(ii)* $\chi(a_0, x) \preceq \chi(a_0, a_1)$ *and* $\chi(x) \preceq \chi(a_1)$, *and*

*(iii)* $\chi(a_0, x) \prec \chi(a_0, a_1)$ *or* $\chi(x) \prec \chi(a_1)$.

*Then $G$ is not finite.*

*Proof.* We choose vertices $a_n$ of $G$ by induction, fulfilling the following properties:

(1) $x, a_0, a_1, \ldots, a_n$ are all distinct,

(2) $i \neq j \Rightarrow \chi(a_i, a_j) \succ 0$,

(3) $\chi(a_0, a_1) \preceq \chi(a_0, a_j)$ and $\chi(a_1) \preceq \chi(a_j)$ for $j \geq 1$,

(4) $\chi(a_j, x) \preceq \chi(a_j, a_k)$ and $\chi(x) \preceq \chi(a_k)$ for $j < k$, and

(5) $\chi(a_0, x) \prec \chi(a_0, a_j)$ or $\chi(x) \prec \chi(a_j)$, for $j \geq 1$.

The basic case is for $n = 1$, and all the clauses here follow from the hypotheses of the lemma.

For the induction step, we assume that $a_0, a_1, \ldots, a_n$ have been chosen fulfilling all the properties so far, and we show how to choose a suitable $a_{n+1}$.

Let $m(n) = \max\{j \in \{1,\dots,n\} : \chi(x,a_j) \succ 0\}$. By (i), $m(n) \geq 1$. Consider the mapping $f_{n+1} = \begin{pmatrix} a_0 & \dots & a_{m(n)-1} & a_{m(n)+1} & \dots & a_n & x \\ a_0 & \dots & a_{m(n)-1} & a_{m(n)+1} & \dots & a_n & a_{m(n)} \end{pmatrix}$, which is a monomorphism from $G[a_0,\dots,a_{m(n)-1},x,a_{m(n)+1},\dots,a_n]$ to $G[a_0,\dots,a_n]$. Since $G$ is **MH**-homogeneous, $f_{n+1}$ has an extension to a homomorphism $f^*_{n+1}$, and we let $a_{n+1} = f^*_{n+1}(a_{m(n)})$, Figure 4.1. We verify properties (1)–(5) for clauses involving the new point $a_{n+1}$.

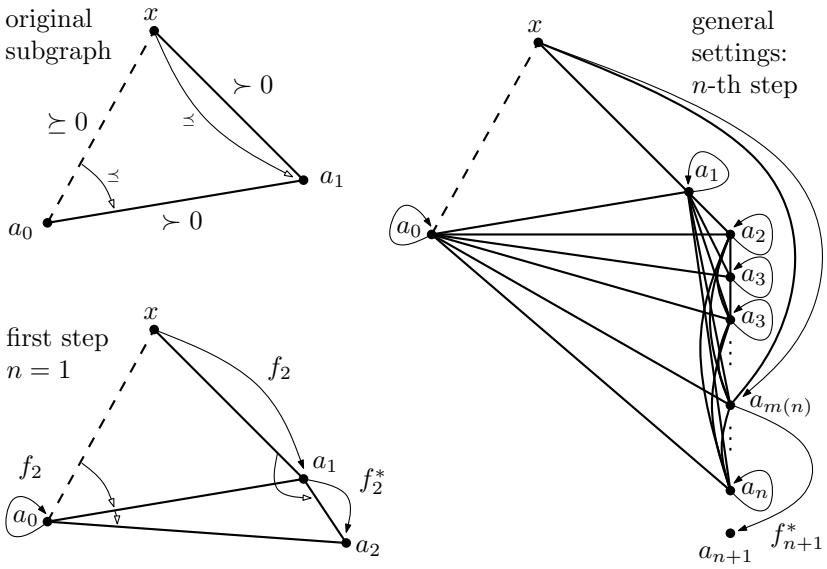



Figure 4.1: The original subgraph, the first step and general setting for the inductive construction. Bold lines without arrows represent edges—solid lines are those having colors $\succ 0$ and dashed those having colors $\succeq 0$. Thin lines with full arrows represents mappings and thin lines with empty arrows indicate the direction of the succession in colors.

Let us show that (1) holds:

- if $a_{n+1} = x$ then $\chi(a_0, a_{m(n)}) \preceq \chi(f^*_{n+1}(a_0), f^*_{n+1}(a_{m(n)})) = \chi(a_0,x)$ and $\chi(a_{m(n)}) \preceq \chi(f^*_{n+1}(a_{m(n)})) = \chi(x)$, which contradicts (5);
- if $a_{n+1} = a_{m(n)}$ then $0 = \chi(a_{m(n)}, a_{n+1}) = \chi(f^*_{n+1}(x), f^*_{n+1}(a_{m(n)})) \succeq \chi(x, a_{m(n)})$, but $\chi(x, a_{m(n)}) \succ 0$ by definition of $m(n)$ – contradiction;
- if $a_{n+1} = a_j$ for some $j \neq m(n)$ then, by (4), $0 \prec \chi(a_j, a_{m(n)}) \preceq \chi(f^*_{n+1}(a_j), f^*_{n+1}(a_{m(n)})) = \chi(a_j, a_{n+1}) = \chi(a_j, a_j) = 0$ – contradiction.

In order to show (2) it suffices to show that $\chi(a_i, a_{n+1}) \succ 0$ for all $i$ such that $0 \leq i \leq n$:

- if $i \neq m(n)$ then $\chi(a_i, a_{n+1}) = \chi(f^*_{n+1}(a_i), f^*_{n+1}(a_{m(n)})) \succeq \chi(a_i, a_{m(n)}) \succ 0$ by the induction hypothesis;
- if $i = m(n)$ then $\chi(a_{m(n)}, a_{n+1}) = \chi(f^*_{n+1}(x), f^*_{n+1}(a_{m(n)})) \succeq \chi(x, a_{m(n)}) \succ 0$ by definition of $m(n)$.

To see that (3) holds we use the induction hypothesis and the fact that $f^*_{n+1}$ is a homomorphism:

- $\chi(a_0, a_1) \preceq \chi(a_0, a_{m(n)}) \preceq \chi(f^*_{n+1}(a_0), f^*_{n+1}(a_{m(n)})) = \chi(a_0, a_{n+1})$;
- $\chi(a_1) \preceq \chi(a_{m(n)}) \preceq \chi(f^*_{n+1}(a_{m(n)})) = \chi(a_{n+1})$.

Let us show (4). As above, from (ii) and (3) we immediately get $\chi(x) \preceq \chi(a_1) \preceq \chi(a_{n+1})$. To see that $\chi(a_j, x) \preceq \chi(a_j, a_{n+1})$ for all $j \in \{0, \ldots, n\}$ we consider several cases:

- if $j > m(n)$ then $\chi(a_j, x) = 0$ by definition of $m(n)$ so $\chi(a_j, x) \preceq \chi(a_j, a_{n+1})$ holds trivially;
- if $j < m(n)$ then using the induction hypothesis and the fact that $f^*_{n+1}$ is a homomorphism we get $\chi(a_j, x) \preceq \chi(a_j, a_{m(n)}) \preceq \chi(f^*_{n+1}(a_j), f^*_{n+1}(a_{m(n)})) = \chi(a_j, a_{n+1})$;
- if $j = m(n)$ then $\chi(a_{m(n)}, x) \preceq \chi(f^*_{n+1}(a_{m(n)}), f^*_{n+1}(x)) = \chi(a_{n+1}, a_{m(n)})$.

Finally, (5) follows from (3) and (iii).

Therefore, $G$ contains an infinite sequence $a_0, a_1, a_2, \ldots$ of pairwise distinct vertices, so it cannot be finite. □

In the rest of the chapter we restrict our attention to two types of partially ordered sets $L$: chains and diamonds.

## 4.2 $L$-colored graphs over chains

In this section we classify finite **MH**-homogeneous $L$-colored graphs where $L$ is a bounded chain and show that in this setting the classes **MH** and **HH** coincide. So, let $L$ be a chain with least element 0 and greatest element 1.

**Lemma 4.2.1.** *Let $G$ be a finite $L$-colored graph which is* **MH***-homogeneous. Assume that $x, y, z$ are three distinct vertices of $G$ satisfying $\chi(x, z) \succ 0$ and $\chi(y, z) \succ 0$. Then:*
*(a) $\chi(x, y) \prec \chi(x, z)$ if and only if $\chi(y) \succ \chi(z)$;*
*(b) $\chi(x, y) = \chi(x, z)$ if and only if $\chi(y) = \chi(z)$.*

*Proof.* Clearly, (b) follows immediately from (a) because $L$ is a chain. Let us show (a). Suppose that $G$ is a finite **MH**-homogeneous $L$-colored graph, and let $x, y, z$ be three distinct vertices of $G$ satisfying $\chi(x, z) \succ 0$ and $\chi(y, z) \succ 0$ but not (a). Then either

$$\chi(x, y) \prec \chi(x, z) \quad \text{and} \quad \chi(y) \preceq \chi(z)$$

or

$$\chi(x, y) \succeq \chi(x, z) \quad \text{and} \quad \chi(y) \succ \chi(z).$$

In both cases finiteness of $G$ contradicts Lemma 4.1.1. □

**Lemma 4.2.2.** *Let $G$ be a finite* **MH***-homogeneous $L$-colored graph. Then:*
*(a) for every $\alpha \in L$, every connected component of $G^{(\alpha)}$ is a uniform graph;*
*(b) for all $x, y \in V(G)$, if $\chi(x, y) \succ 0$ then $\chi(x) = \chi(y)$;*
*(c) every connected component of $G$ is a uniform graph.*

*Proof.* (*a*) Take any $\alpha \in L$ and let $S$ be a connected component of $G^{(\alpha)}$. Then, by the definition of $G^{(\alpha)}$, we have that $\chi(x) = \alpha$ for all $x \in S$. Let us show that $\chi(x, y)$ is constant for all $x, y \in S$ satisfying $x \neq y$. If $|S| = 1$ or $|S| = 2$ the claim is trivial. Assume that $|S| \geq 3$. Since $S$ is a connected component, it suffices to show that whenever $x, y, z \in S$ are three distinct vertices such that $\chi(x, z) \succ 0$ and $\chi(y, z) \succ 0$, then $\chi(x, z) = \chi(y, z) = \chi(x, y)$. So, let $x, y, z \in S$ be three distinct vertices satisfying $\chi(x, z) \succ 0$ and $\chi(y, z) \succ 0$. Since $\chi(y) = \chi(z) = \alpha$, Lemma 4.2.1 yields that $\chi(x, y) = \chi(x, z)$. Analogously, $\chi(x, y) = \chi(y, z)$.

(*b*) Assume that there exist $x_1, x_2 \in V(G)$ such that $\chi(x_1, x_2) \succ 0$ and $\chi(x_1) \neq \chi(x_2)$. Without loss of generality we can assume that $\chi(x_1) \prec \chi(x_2)$. Let us now construct a sequence of vertices $x_3$, $x_4$, ... inductively so that $\chi(x_{i-1}) \preceq \chi(x_i)$ for all $i \geq 2$. Assuming that $x_n$ has been chosen, since $\chi(x_{n-1}) \preceq \chi(x_n)$, the mapping $f_{n+1}$ taking $x_{n-1}$ to $x_n$ is a monomorphism from $G[x_{n-1}]$ to $G[x_n]$, so as $G$ is **MH**-homogeneous, there is an endomorphism $f^*_{n+1}$ of $G$ which extends $f_{n+1}$ and we let $x_{n+1} = f^*_{n+1}(x_n)$. Then $\chi(x_n) \preceq \chi(f^*_{n+1}(x_n)) = \chi(x_{n+1})$ giving the induction step. Note, also, that

$$\chi(x_{i-1}, x_i) = \chi(f^*_i(x_{i-2}), f^*_i(x_{i-1})) \succeq \chi(x_{i-2}, x_{i-1}), \text{ for all } i.$$

Therefore, we have constructed a sequence of vertices $x_1, x_2, x_3, \ldots$ such that $\chi(x_1) \preceq \chi(x_2) \preceq \chi(x_3) \preceq \ldots$ and $0 \prec \chi(x_1, x_2) \preceq \chi(x_2, x_3) \preceq \chi(x_3, x_4) \preceq \ldots$. Since $\chi(x_1) \prec \chi(x_2)$ and since $G$ is finite there exists an $n$ such that $\chi(x_{n-2}) \prec \chi(x_{n-1}) = \chi(x_n)$. Then Lemma 4.2.1 yields that $\chi(x_{n-2}, x_{n-1}) = \chi(x_{n-2}, x_n) \succ 0$ since $\chi(x_{n-1}) = \chi(x_n)$. By the same lemma we also have $\chi(x_{n-1}, x_n) \prec \chi(x_{n-2}, x_{n-1})$ since $\chi(x_n) \succ \chi(x_{n-2})$ . On the other hand, $\chi(x_{n-1}, x_n) \succeq \chi(x_{n-2}, x_{n-1})$ by construction. Contradiction.

(*c*) It follows from (*b*) that $S$ is a connected component of $G$ if and only if $S$ is a connected component of $G^{(\alpha)}$ for some $\alpha \in L$. Therefore, every connected component of $G$ is a uniform graph. □

**Theorem 4.2.3.** *Let $G$ be a finite $L$-colored graph where $L$ is a chain with the least element 0 and the greatest element 1. Then the following are equivalent:*

*(1) $G$ is* **HH***-homogeneous,*

*(2) $G$ is* **MH***-homogeneous,*

*(3) $G$ has the following structure:*

- *every connected component of $G$ is a uniform $L$-colored graph, and*
- *if $U(n_1, \alpha_1, \beta_1)$ and $U(n_2, \alpha_2, \beta_2)$ are connected components of $G$ such that $\alpha_1 \preceq \alpha_2$, then $n_1 \leq n_2$ and $\beta_1 \preceq \beta_2$. Consequently, if $\alpha_1 = \alpha_2$, then $n_1 = n_2$ and $\beta_1 = \beta_2$.*

*Proof.* (3) ⇒ (1) is easy.

(1) ⇒ (2) is obvious.

(2) ⇒ (3). Let $G$ be a finite **MH**-homogeneous $L$-colored graph. We already know from Lemma 4.2.2 that every connected component of $G$ is a uniform graph. So, let $S_1$ and $S_2$ be connected components of $G$ such that $G[S_1] \cong U(n_1, \alpha_1, \beta_1)$, $G[S_2] \cong U(n_2, \alpha_2, \beta_2)$ and assume that $\alpha_1 \preceq \alpha_2$. Let $x$ be an arbitrary vertex of $S_1$ and $y$ an arbitrary vertex of $S_2$. Then $f$ taking $x$ to $y$ is a monomorphism

from $G[x]$ to $G[y]$, since $\chi(x) = \alpha_1 \preceq \alpha_2 = \chi(y)$. So, by the homogeneity requirement, $f$ extends to an endomorphism $f^*$ of $G$. It is easy to see that an endomorphism maps a connected component of $G$ into another connected component of $G$, so $f^*(S_1) \subseteq S_2$, since $f^*(x) = y \in S_2$. Moreover, $f^*|_{S_1}$ is injective (assume that $x, y \in S_1$ are two distinct vertices such that $f^*(x) = f^*(y)$; then $\chi(f^*(x), f^*(y)) = 0$ because $G$ is without loops; on the other hand, $\chi(f^*(x), f^*(y)) \succeq \chi(x, y) = \beta_1 \succ 0$ by the definition of an edge-uniform $L$-colored graph – contradiction), so $n_1 = |S_1| \leq |S_2| = n_2$. Finally, if $x, y \in S_1$ are two distinct vertices, then $\beta_1 = \chi(x, y) \preceq \chi(f^*(x), f^*(y)) = \beta_2$. □

## 4.3 $L$-colored graphs over diamonds

In this section we consider $L$-colored graphs where $L$ is a diamond. We first consider finite vertex-uniform $L$-colored graphs and show that in this case the classes **MH** and **HH** coincide. We then provide an example of an $L$-colored graph which is **MH**-homogeneous, but not **HH**-homogeneous, proving thus that in the general case the classes **MH** and **HH** do not coincide for $L$-colored graphs when $L$ is a diamond. So, let $L$ be a diamond with least element 0 and greatest element 1.

First, we consider finite **MH**-homogeneous vertex-uniform $L$-colored graphs such that $\chi(x, y) = 1$ for some pair of distinct vertices $x$ and $y$. At this point it is important to stress that Lemma 4.3.1 as well as Propositions 4.3.2 and 4.3.3 hold for arbitrary partial orders $L$, not only diamonds. However, these three statements are important prerequisites for Theorems 4.3.4 and 4.3.5 which essentially depend on the requirement that $L$ be a diamond.

**Lemma 4.3.1.** *Let $G$ be a finite* **MH***-homogeneous vertex-uniform $L$-colored graph and assume that there exist $x_0, y_0 \in V(G)$ such that $\chi(x_0, y_0) = 1$. Then the following holds:*

*(1) For every vertex $x$ there is a vertex $y$ such that $\chi(x, y) = 1$.*

*(2) Let $x, y, z$ be distinct vertices. If $\chi(x, y) = \chi(y, z) = 1$ then $\chi(x, z) = 1$.*

*(3) If $x$ and $y$ belong to the same connected component of $G$ then $\chi(x, y) = 1$.*

*Proof.* (1) Let $x$ be an arbitrary vertex. Then $f$ taking $x_0$ to $x$ extends to an endomorphism $f^*$ of $G$, so $\chi(x, f^*(y_0)) = \chi(x_0, y_0) = 1$.

(2) Let $\chi(x, y) = \chi(y, z) = 1$. If $\chi(x, z) \prec 1$, Lemma 4.1.1 yields that $G$ then cannot be finite. Contradiction.

(3) Let $S$ be a maximal set of vertices of $G$ such that $x \in S$ and $\chi(u, v) = 1$ for all $u, v \in S$ with $u \neq v$. Note that $|S| \geq 2$ due to (1). Let us show that $S$ coincides with the connected component $W$ of $G$ that contains $x$. Suppose to the contrary that this is not the case and take any $z \in W \setminus S$ such that $\chi(z, y) \succ 0$ for some $y \in S$. If $\chi(z, s) = 1$ for some $s \in S$ then, by(2), $\chi(z, s) = 1$ for all $s \in S$, which contradicts the choice of $z$. Therefore, $\chi(z, s) \prec 1$ for all $s \in S$. Take any $w \in S \setminus \{y\}$. As we have just concluded, $0 \prec \chi(z, y) \prec 1$ and $0 \preceq \chi(z, w) \prec 1$, so Lemma 4.1.1 yields that $G$ is not finite. Contradiction. □

As an immediate corollary, we get the following straightforward result:

**Proposition 4.3.2.** *Let $G$ be a finite* **MH***-homogeneous vertex-uniform $L$-colored graph where every vertex has color $\alpha \in L$. Assume that there exist $x_0, y_0 \in V(G)$ such that $\chi(x_0, y_0) = 1$. Then there exists a positive integer $n$ such that every connected component of $G$ is isomorphic to $U(n, \alpha, 1)$.*

*Proof.* By Lemma 4.3.1, every connected component $S_i$ of $G$ is of the form $U(|S_i|, \alpha, 1)$. If $|S_i| > |S_j|$ for some distinct connected components $S_i$ and $S_j$ of $G$, then a partial monomorphism which maps a vertex from $S_i$ to a vertex from $S_j$ would, by the homogeneity requirement, extend to an endomorphism of $G$ which would then map $S_i$ into $S_j$. On the other hand, such an endomorphism would necessarily be injective on $S_i$ because of our requirement that $\chi(x, x) = 0$ (see Section 4.1). Contradiction. □

Next, we consider finite **MH**-homogeneous vertex-uniform $L$-colored graphs satisfying $\chi(x, y) = 1$ for no $x, y \in V(G)$.

**Proposition 4.3.3.** *Let $G$ be a finite connected* **MH***-homogeneous vertex-uniform $L$-colored graph such that $\chi(x, y) = 1$ for no $x, y \in V(G)$. Then $G$ is complete.*

*Proof.* Assume, to the contrary, that $G$ is not complete. Then there exist $x, y \in V(G)$ such that $x \neq y$ and $\chi(x, y) = 0$. Since $G$ is connected, there exists a sequence $v_1, v_2, \ldots, v_k$ of vertices of $G$ such that $x = v_1$, $y = v_k$ and $\chi(v_i, v_{i+1}) \succ 0$ for all $i \in \{1, \ldots, k-1\}$. Without loss of generality, we can assume that $(v_1, v_2, \ldots, v_k)$ is the shortest such sequence, so that $\chi(v_i, v_j) = 0$ whenever $j - i > 1$. Note that $k \geq 3$ beacuse $\chi(x, y) = 0$. Now, $v_1$, $v_2$ and $v_3$ provide a configuration which, by Lemma 4.1.1, ensures that $G$ is not finite. Contradiction. □

If $G$ is a finite vertex-uniform $L$-colored graph which is connected and complete, all endomorphisms are automorphisms, and it is easy to see that $G$ is **HH**-homogeneous if and only if $G$ is **MH**-homogeneous if and only if $G$ is ultrahomogeneous. On the other hand, if $G$ is a finite vertex-uniform $L$-colored graph wich is not connected and has the property that $\chi(x, y) \prec 1$ for all $x, y \in V(G)$, then by Proposition 4.3.3 every connected component of $G$ is complete and all components have to be be isomorphic. So, we have the following partial classification result which depends on the classification of all finite ultrahomogeneous edge-colored graphs (and this is a long-standing open problem):

**Theorem 4.3.4.** *Let $L$ be a diamond with least element 0 and greatest element 1. The following are equivalent for a finite vertex-uniform $L$-colored graph $G$ where every vertex is colored by $\alpha \in L$:*

*(1) $G$ is* **HH***-homogeneous,*

*(2) $G$ is* **MH***-homogeneous,*

*(3) $G$ is a disjoint union of $k \geq 1$ copies of $H$, where*

- *$H$ is $U(n, \alpha, 1)$ for some positive integer $n$; or*
- *$H$ is an ultrahomogeneous $L$-colored graph such that $0 \prec \chi(x, y) \prec 1$ for all distinct $x, y \in V(G)$, and $\chi(x) = \alpha$ for all $x \in V(G)$.*

*Proof.* The implications (3) $\Rightarrow$ (1) $\Rightarrow$ (2) are straightforward. Let us show (2) $\Rightarrow$ (3). If $\chi(x,y) = 1$ for some distinct $x, y \in V(G)$, then Proposition 4.3.2 yields that there is a positive integer $n$ such that every connected component of $G$ is isomorphic to $U(n, \alpha, 1)$. Assume, therefore, that $\chi(x,y) \prec 1$ for all $x, y \in V(G)$. By Proposition 4.3.3 every connected component of $G$ is complete. Therefore, for all distinct $x$ and $y$ in the same connected component we have that $0 \prec \chi(x,y) \prec 1$. Since $L$ is a diamond, it easily follows that for every endomorphsim $f$ and every pair of vertices $x$ and $y$ in the same connected component we have $\chi(x,y) = \chi(f(x), f(y))$. From this we then easily infer that every connected component of $G$ is an ultrahomogeneous $L$-colored graph and that all connected components of $G$ are isomorphic. □

In particular, if $L = M_2$ is the diamond on four elements $0, b, r, 1$ where $0 \prec b \prec 1$, $0 \prec r \prec 1$ and where $b$ and $r$ are incomparable ($b$ and $r$ stand for *blue* and *red*, respectively), we can provide the full classification as follows. For an $\alpha \in M_2$ let $G_{(\alpha)} = (V, E_\alpha)$ be the (ordinary undirected) graph where $E_\alpha = \{\{x,y\} : \chi(x,y) = \alpha\}$.

**Theorem 4.3.5.** *The following are equivalent for a finite vertex-uniform $M_2$-colored graph $G$ where every vertex is colored by $\alpha \in M_2$:*

*(1) $G$ is* **HH***-homogeneous,*

*(2) $G$ is* **MH***-homogeneous,*

*(3) $G$ is a disjoint union of $k \geq 1$ copies of $H$, where*

- *$H$ is $U(n, \alpha, 1)$ for some positive integer $n$; or*
- *$H$ is vertex uniform, $H_{(r)}$ is one of the Gardiner graphs and $H_{(b)}$ is its complement.*

Note that part of this theorem is just a reformulation of Theorem 3.2.5. It additionally provides an argument for the coincidence of classes **HH** and **MH** for finite vertex-uniform $M_2$-colored graphs, effectively proving Proposition 3.3.1.

As Example 4.3.1 below shows, Theorems 4.3.4 and 4.3.5 cannot be extended to finite $L$-colored graphs when $L$ is a diamond and graphs are not required to be vertex-uniform. In fact for both cases the classes coincide when the corresponding $L$-colored graphs are vertex-uniform.

**Example 4.3.1.** Let $G$ be an $M_2$-colored graph on four vertices $a, b, c, d$ where the vertices and the edges are colored as follows: $\chi(a) = \chi(b) = r$, $\chi(c) = \chi(d) = b$, $\chi(a,c) = \chi(c,d) = \chi(b,d) = r$, $\chi(a,d) = \chi(b,c) = b$ and $\chi(a,b) = 0$ (see Figure 4.2).

Then $G$ is clearly an **MH**-homogeneous graph. To see that $G$ is not an **HH**-homogeneous graph it suffices to note that the partial homomorphism $f = \binom{a\ b}{a\ a}$ cannot be extended to an endomorphism of $G$.

This example enables us to state the following proposition that $L$-colored graphs are in fact sufficient for distinguishing the classes **MH** and **HH**.

**Proposition 4.3.6.** *There is a partial order $L$ and corresponding $L$-colored graphs for which the classes* **MH** *and* **HH** *do not coincide.*

. □

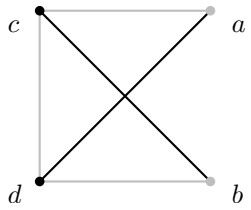



Figure 4.2: An example of a finite $L$-colored graph that is **MH**-homogeneous but not **HH**-homogeneous.

## 4.4 Concluding remarks

All results presented in this chapter, except several new propositions added by the author, can be found in two consecutive papers [61, 57].

The simple relational structure presented in Example 4.3.1 can easily be generalised to provide a whole class of structures that are all **MH**-homogeneous but not **HH**-homogeneous.

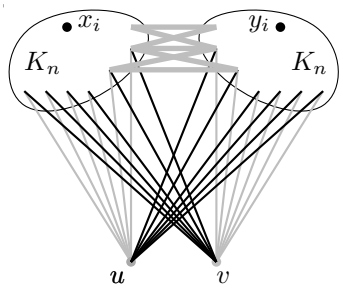



Figure 4.3: A class of finite $L$-colored graphs which are all **MH**- but not **HH**-homogeneous. Cliques $K_n$ are colored black on vertices and edges. The cliques are connected by thick gray lines, the two vertices $u$ and $v$ colored gray are joined by black edges to one clique and by gray edges to the other clique.

The construction is depicted in Figure 4.3. Fix $n \in \{1, 2, \ldots, \omega\}$. (Note that in case $n = \omega$ we get an example of a countably infinite structure that is **MH**-homogeneous but not **HH**-homogeneous.) Take two cliques both of size $n$ whose vertices and edges are colored black. Join the vertices of these two cliques by thick gray edges. Finally, add two new nonadjacent vertices $u$ and $v$ colored gray, and join the two vertices and the vertices of the two cliques by black and gray edges as in Figure 4.3. Then, as in Example 4.3.1, we can show that this graph is **MH**-homogeneous but not **HH**-homogeneous.

A question that immediately arises is whether one can avoid the need for colored vertices at the expense of introducing loops. Consider the finite edge-colored graph depicted in Figure 4.4 with no colors assigned to vertices that we construct as follows. Given $n > 1$, take five copies of $K_n$ and color their edges gray. Now join these cliques by complete bipartite graphs using two mutually disjoint 5-cycles where the edges of one 5-cycle are black, while the edges of the other 5-cycle are gray. Furthermore, add a black loop to each vertex.

This graph is easily seen to be **MH**-homogeneous. To see that it is not **HH**-homogeneous, consider a partial homomorphism unifying two neighboring cliques

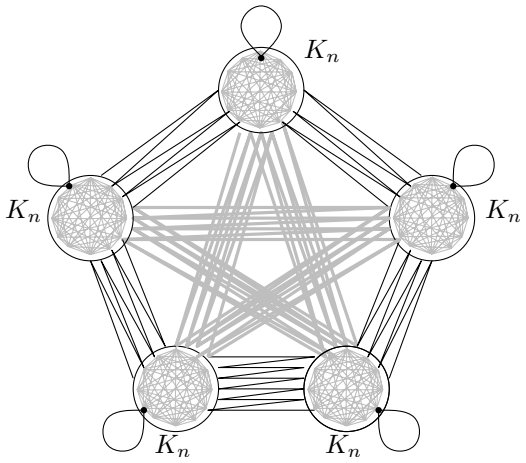



Figure 4.4: An edge-colored graph with loops that is **MH**- but not **HH**-homogeneous.

(this is possible due to the black-colored loops). Then every endomorphism that extends such a partial homomorphism would enforce the existence of an edge colored both black and gray.

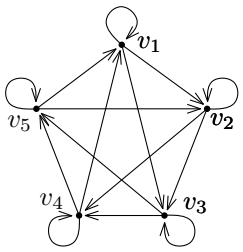



Figure 4.5: A digraph with loops that is **MH**-homogeneous but not **HH**-homogeneous.

Finally in Figure 4.5 we present a directed graph with loops that is **MH**-homogeneous but not **HH**-homogeneous. To see that this digraph is not **HH**-homogeneous consider a partial homomorphism $f = \begin{pmatrix} v_1 & v_4 & v_5 \\ v_1 & v_5 & v_5 \end{pmatrix}$. Then every endomorphism that extends $f$ would enforce the existence of a bidirectional edge.

A natural generalization of the proposed classifications is to use a less restrictive collection of colors. This leads to following problem.

**Problem 4.4.1.** *Let $L$ be an arbitrary partially ordered set. Classify all finite* **HH***-homogeneous and* **MH***-homogeneous $L$-colored graphs.*

Solving this problem could be hard, although only finite structures are considered.

Proposition 3.3.1 and Proposition 4.3.6 prompt the question of analysing the coincidence of classes **MH** and **HH**. This analysis should lead to the identification of a minimal extension of ordinary graphs that would distinguish the classes **MH** and **HH**. It seems that vertex-uniformity is the key property in this regard.

**Conjecture 4.4.1.** Do the classes **MH** and **HH** coincide for finite vertex-uniform $L$-colored graphs?

Once these problems are solved we could follow their trail into the countably infinite world, where examples of homomorphism-homogeneous structures could be even more interesting. Since posing the problem of classifying countable homomorphism-homogeneous $L$-colored graphs would be premature, it is more appropriate to deal with distinguishing the classes **MH** and **HH**. Assuming Conjecture 4.4.1 to be true, we can formulate the following problem that extends it.

**Problem 4.4.2.** *Do the classes* **MH** *and* **HH** *coincide for countable vertex-uniform $L$-colored graphs?*

# 5. Morphism extension classes

Stones grow, plants grow and live,
animals grow, live and feel.

Carl Linnaeus

This chapter summarizes results dealing with various homogeneity classes, like the classes **MH** and **HH** studied in Chapters 3 and 4 – see in particular the concluding remarks in Sections 3.3 and 4.4. More precisely it discusses properties of relations between the classes of structures defined by the various types of homogeneity of Definition 1.7.1. Inspired by the distinction between the classes **HH** and **MH** for $L$-colored graphs given by Proposition 4.3.6, we study mutual coincidence between other pairs of classes.

Similarly to the case of the **HH** and **MM** correspondence it is possible to ask for the least "complex" structure that distinguishes a pair of classes from each other. Here, by complexity we, in vague terms, mean the richness of the corresponding relational structure definition, for example the number of relations added. The main purpose of this chapter is to summarize the results of the previous chapter and some from the literature and to put them in a broader context of the classes given by Definition 1.7.1, called here *morphism extension classes*.

## 5.1 Known hierarchies

It is easily observed that the class of finite **HH**-homogeneous undirected graphs is significantly smaller compared to the ultrahomogeneous case. More precisely, for finite graphs there are only complete and null graphs in the class **HH** [18]. Such an inclusion is to be found in many cases. For example **MH** is a subclass of **IH**. The obvious inclusions between the morphism extension classes are depicted in Figure 5.1. Note that, for simplicity, we omit the inclusions implied by transitivity in all diagrams and following [17] we omit the lower part of the hierarchy containing the extension of less strict morphisms into morphisms that are more strict like the class **HI**.

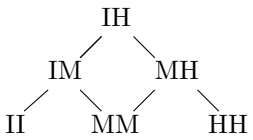
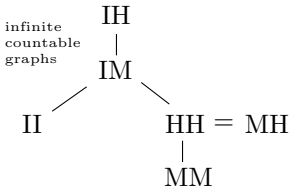


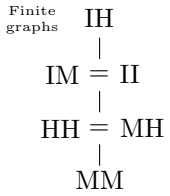

Figure 5.1: The hierarchy of morphism extension classes for a general relational structure. A line connecting two classes means that the class horizontally below the other is contained in it. For simplicity, all lines implied by transitivity are omitted.

If one fixes the type of structure, the hierarchy often differs from the general one of Figure 5.1, usually by identifying particular classes. A good example

is the hierarchy of extension morphism classes for graphs due to Rusinov and Schweitzer [136] depicted in Figure 5.2.




(a) Countably infinite graphs

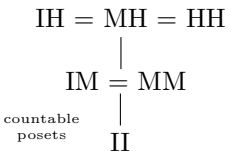



(b) Finite graphs

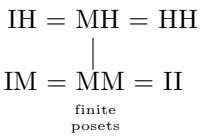
Figure 5.2: The hierarchy of morphism extension classes for graphs. The meaning of the lines is the same as in Figure 5.1. Adapted from [136].

Here we can see the already discussed coincidence of the classes **MH** and **HH** for countable graphs. Adapting results from Rusinov and Schweitzer [136], we can check that the class **IM** is a proper subclass of **IH**, which can be shown using $K_{n,1}$ for finite graphs and $K_{\omega,1}$ for countably infinite graphs, respectively. On the other hand, going from the finite case to the countably infinite case the classes **II** and **IM** start to be different. One witness of this difference is a countably infinite graph created from the infinite clique by removing one edge.

Another example is given by partially ordered sets with nonstrict order, for which the hierarchy is as depicted in Figure 5.3, adapted from [17]. For a deeper description of the situation that holds here see [17].




(a) Countably infinite posets




(b) Finite posets

Figure 5.3: Hierarchy of morphism extension classes for partially ordered sets [17]. The meaning of the lines is the same as in Figure 5.1. Adapted from [17]

The usual task of classification is to provide structures or their parametric classes that are homogeneous in the sense of the corresponding morphism extension class. One of the questions around these morphism extension classes is the following problem – again by structural complexity here we mean a vaguely defined notion that includes the number and character of defined relations.

**Problem 5.1.1.** *How far, in the sense of underlying structural complexity, does one have to go in order to make all the XY-homogeneous classes distinct, where* $X, Y \in \{\mathbf{H}, \mathbf{M}, \mathbf{I}\}$*?*

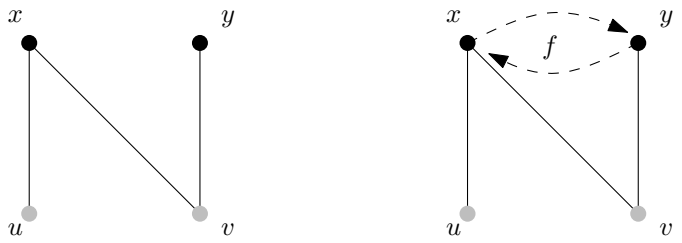
This has been partially discussed in Chapter 4, where the question about the relationship between **HH** and **MH** was solved for $L$-colored graphs. In all previous existing classification results the classes **HH** and **MH** have coincided.

## 5.2 Classes for $L$-colored graphs

To show that the class **IH** is different from the classes **IM** and **MH** we can use the results for the graphs mentioned in Section 5.1. The star graph, i.e. the complete bipartite graph with partites $V_1$ of size 1 and $V_2$ of size $n-1$, is an obvious example of a graph that is **IH**, but neither **IM** nor **MH**. Its containment in the class **IH** can be simply seen by considering two basic types of nontrivial local isomorphism:

1. permutations on $V_2$, and
2. isomorphisms mapping a vertex $v \in V_1$ to some vertex $u \in V_2$ and another vertex $u' \in V_2$, not necessarily different from $u$, to vertex $v$.

Both these local isomorphisms can be easily extended to an endomorphism. On the other hand, a star cannot be **IM**, while the second-mentioned local isomorphisms cannot be extended to a global monomorphism over the whole graph. At the same time a star graph cannot be **MH**, while a monomorphism fixing any vertex $u \in V_2$ and mapping another vertex $u' \in V_2$ such that $u \neq u'$ to vertex $v \in V_1$ cannot be extended to a homomorphism – consider the possible images for vertex the $v \in V_1$. This shows that the class **IH** is distinct from all others in the case of graphs.




(a) **MH** vertex-colored graph (b) Local isomorphism $f$ which shows that $MH \neq IM, MM, II$

Figure 5.4: **MH** vertex colored graph that is neither **IM** nor **MM** nor **II**.

To show a similar property for the class **MH** one can make use of the vertex-colored graph $H$ created from the path $P_3$ of length 3 with an alternating coloring of its vertices, i.e. for $V(H) = V(P_3) = (u, x, v, y)$ we color vertices $x$ and $y$ by the first color and vertices $u$ and $v$ by the second, as depicted in Figure 5.4(a). This graph is obviously **MH**. To show that it is neither **IM** nor **MM** nor **II** one can use a local isomorphism $f$ defined by $f(x) = y$ and $f(y) = x$ which cannot be extended to any monomorphism – see Figure 5.4(b). This graph can be also used to distinquish the class **HH** from the classes **MM**, **IM** and **II**, since the graph $H$ is obviously **HH**-homogenous but neither **MM** nor **IM** nor **II**, as shown above.

For a pair of classes **MH** and **HH** we can use Example 4.3.1 from Chapter 4 giving an $L$-colored graph which witnesses the fact that the classes **MH** and **HH** are distinct. This example is however even more valuable. It can also be used to distinguish the classes **HH** and **MM** and similarly **II**– note that the graph $G$ from the example is obviously **MM** and **II**.

To distinguish the class **IM** from the others, a simple homogeneous graph, namely the cycle $C_5$ of length 5, can be used. Homogeneity of this graph makes it a member of **IM**. Let $V(C_5) = \{v_1, v_2, \ldots, v_5\}$ and $E(C_5) = E_1 \cup \{v_5, v_1\}$ where $E_1 = \{\{v_i, v_{i+1}\}; i = 1, 2, 3, 4\}$. Then a local monomorphism $h$ defined by $h(v_1) = v_1$ and $h(v_3) = v_2$ cannot be extended to an endomorphism. Since every monomorphism is a homomorphism this graph is neither **MM** nor **MH** nor **HH**. In the same way $C_5$ is seen to be a graph that is **II**, but neither **MH** nor **MM** nor **HH**.

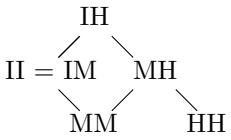



Figure 5.5: Morphism extension classes hierarchy for finite vertex- and edge-colored graphs.

Finally, considering only finite $L$-colored graphs one can show that classes **II** and **IM** coincide. This can be seen from the simple fact that in a finite graph every surjective monomorphism is in fact an isomorphism. This can be shown by a simple counting argument starting with an intial map from a non-edge to an edge and then a consequent series of mappings of corresponding edges. Since the graph is finite this shows that such an initial mapping is impossible.

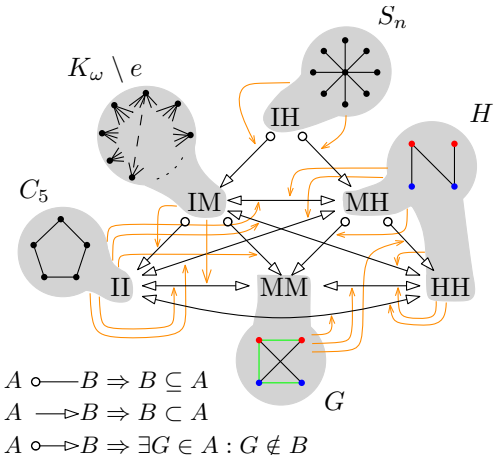



Figure 5.6: General reasoning when constructing the complete hierarchy for countable $L$-colored graphs. Gray areas indicate important participation for the corresponding graph – some of the classes in which the corresponding graph belongs are not depicted in order to preserve readability of the figure. Black lines indicate relationships between pairs of classes. Some of the lines are also omitted – mostly those that can be simply derived from the above-mentioned results. The particular meaning of arrows are shown in the figure itself. On the other hand, orange lines (see as gray in black and white print) indicate which relationships can be explained by the corresponding graph.

Summarizing all the previously mentioned results it is possible to show a final form for the hierarchy of morphism extension classes for finite $L$-colored graphs

as depicted in Figure 5.5. This diagram however also provides reasoning about morphism extension classes for infinite $L$-colored graphs – for more details see below.

As a final observation, for finite graphs the prospects do not seem very promising for distinguishing the classes **II** and **IM**. On the other hand, when considering countable graphs one can easily find a graph that is **IM** but not **II**. For this purpose a graph from [136] can be used. It is an infinite clique with one edge removed, i.e. $K_\omega \setminus e$. This graph is not in the class **II**, since a local isomorphism mapping a vertex that is incident to the removed edge to another vertex that is not incident to this edge cannot be extended by isomorphism – consider the image of the second vertex of the non-edge. On the other hand it is in the class **IM** since in this case this non-edge can be mapped to an edge. At the same time the previously mentioned counting argument is not a problem since we have an infinite clique. By a simple extension of this argument it can be shown that this graph is **MM**, thus distinguishing the classes **II** and **MM**. For this reason the hierarchy of morphism extension classes for countable $L$-colored graphs has the general structure as shown in Figure 5.1. General reasoning for the whole process can be checked in Figure 5.6.

## 5.3 Concluding remarks

This chapter contains virtually nothing new, except several examples of coloured graphs distinguishing some classes and a few references to other sections, like Example 4.3.1. Most of the mentioned properties are either well known or derived from the papers [136, 109, 17]. The purpose of this chapter is to summarize these results from the perspective of morphism extension classes and to describe the whole hierarchy, which in turn partially answers Problem 5.1.1.

There are several possibilities for extending such an analysis. Several open problems dealing mainly with the classes **MH** and **HH** are mentioned in Chapters 3 and 4. Recently, there have been some preliminary results dealing with a variant of homogeneity called connected-homomorphism-homogeneity – see Section 1.9 and the paper [97]. In the paper just cited several new classes for this type of homomorphism-homogeneity are given. This leads to the following:

**Problem 5.3.1.** *What is the morphism extension classes hierarchy for $L$-colored graphs when using connected-homomorphism-homogeneity?*

Another future direction for research could be in using other recent preliminary findings introducing new morphism extension classes for specific types of morphism – see [98]. Roughly speaking, this work introduces another two morphisms into Definition 1.7.1. These are epimorphism and bimorphism. An *epimorphism* is just a surjective endomorphism and a bimorphism is epimorphism which is at the same time a monomorphism. Another step could thus be to explore these new classes for colored structures.

# 6. Relational complexity

> Everything is both simpler than
> we can imagine, and more
> complicated that we can conceive.
>
> Johann Wolfgang von Goethe

The results of previous chapters suggest that there are not many structures that are homogeneous, even under variant definitions of homogeneity. For instance, staying in the world of ultrahomogeneous structures, there are very few examples among undirected graphs. This situation may change when the underlying language for structures is extended. Theorem 3.2.5 provides the list homomorphism-homogeneous bicolored graphs. It can be easily seen that these structures are also ultrahomogeneous, and thus all homomorphisms are in fact also isomorphisms. This shows that Gardiner's ultrahomogeneous graphs given in Theorem 1.4.2 can be extended by introducing colors on edges, even though this extension is quite simple. Moreover, it suggests that this process can be somehow generalized. For the sake of brevity, we only discuss ultrahomogeneity here (unless otherwise stated).

Having a classification of ultrahomogeneous relational structures in hand for a given language $L$, we can extend this language to a language $L'$ and study the classification for this extension. Since all ultrahomogeneous structures in the language $L$ can also be realized in the language $L'$, the classification for this extension can only be wider. We could follow the pattern of our previous efforts and look for new structures while introducing new relations. However this process can be also converted, prompting the following roughly stated question.

**Problem 6.0.2.** *For a given relational structure* $\mathbf{A}$*, what is the minimal arity* $r$ *such that there exists a set of relations* $S$ *each of arity at most* $r$ *such that the structure* $\mathbf{A}'$ *constructed as an expansion of* $\mathbf{A}$ *using relations from* $S$ *is ultrahomogeneous.*

This question should be given careful interpretation and underlying assumptions made clear. Taking the roughly defined terms of the question as it stands, it is always possible to make a finite structure ultrahomogeneous by coloring each vertex differently. This immediately trivializes our question for finite structures. Nevertheless, as it stands the it can still be interesting for infinite structures since we can adopt the natural assumption that the number of relations is finite, at least for each arity. This is discussed in more detail in this chapter and leads to the notion of lift complexity.

Still, it remains to be decided whether our question for finite structures can be interesting from some point of view. The extension just described that gives vertices different colors violates the desirable property of extensions that they do not affect the automorphism group of the structure– coloring all the vertices with different colors makes the automorphism group of the underlying structure contain only the identity, and moreover also forces all local isomorphisms to be trivial. Since the criterion for a structure to be ultrahomogeneous involves the extension of local isomorphisms to automorphisms, such an extension that reduces

all local isomorphisms to trival ones should be considered as non-preserving. On the other hand, such symmetry-breaking extensions remain interesting for infinite structures, as mentioned above.

Adding the condition that an extension has to "not harm the automorphism group" results in the notion of relational complexity defined below. Our revised question now reads as follows:

**Problem 6.0.3.** *For a given relational structure* $\mathbf{A}$ *what is the minimum arity* $r$ *such that there exists a set of relations* $S$ *each of arity at most* $r$ *such that the structure* $\mathbf{A}'$ *constructed as an expansion of* $\mathbf{A}$ *using relations from* $S$ *is ultrahomogeneous while at the same time* $\mathrm{Aut}(\mathbf{A}) = \mathrm{Aut}(\mathbf{A}')$.

In Section 6.5 of this chapter both these questions are posed not only about a single relational structure, but also about universal structures for a class of structures. In this case bounds for arities rather than arities themselves are studied, especially because there can be many universal structures for a given class.

This chapter is based on two consecutive results [59, 60]. Examples are given within the flow of the chapter and results are always referenced at their point of use. This work is based on a slightly different notion defined for permutation groups that appeared in [29, 27]. For a more detailed discussion see Section 6.3.1.

In Section 6.2 these concepts are formalized as complexity characteristics and discussed for various classes of structures and complexity values. First to be considered is the finite case for ordinary undirected graphs. This is discussed in Section 6.3, where small values of complexities that are already interesting are analysed, namely in Subsections 6.3.1 and 6.3.2. On the other hand, high values of relational complexity for graphs are discussed in Section 6.4. The chapter closes with Section 6.5, which contains results on bounds for both complexities in the case of infinite relational structures. This last section draws its inspiration mainly from [73, 32].

## 6.1 Introduction

The classification of ultrahomogeneous undirected graphs given by Theorem 1.4.2 shows that cycles are ultrahomogeneous up to size 5 ($C_3$ is a complete graph, $C_4$ is a complete bipartite graph, $C_5$ is a sporadic case). Similarly to the Petersen graph, $C_6$ has two different independent sets, only here of size 2 (a pair of vertices at distance 2 and a pair of vertices at distance 3). $C_6$ can be turned into an ultrahomogeneous structure by introducing another type of edge, say red, and by connecting every pair of vertices at distance 2 using these newly defined red edges. The resulting structure is ultrahomogeneous and we consider such an extended structure to be a *homogenization* of the original structure. Cherlin [22], when explaining a more general theory of ultrahomogeneous structures [86], notices that such a process can be seen more generally — classifications of ultrahomogeneous structures usually contain sporadic ultrahomogeneous structures that can be considered as members of regular ultrahomogeneous families in a suitably extended language. Although $C_5$ is sporadic as an ultrahomogeneous graph, it belongs to the family of graph cycles all of which are ultrahomogeneous as metric graphs [23].

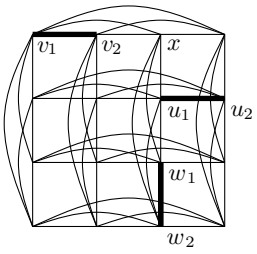



Figure 6.1: $L(K_{n,n})$: graph containing two subgraphs, each isomorphic to $2 \cdot K_2$ (two disjoint edges), with the property that any isomorphism between them cannot be extended to an automorphism of whole graph. This example was given originally in [22].

Just as for cycles, one can observe the extension of a sporadic case into a whole class of ultrahomogeneous structures when introducing new relations for the graph $L(K_{3,3})$. What is so special about the structure of this graph that makes it ultrahomogeneous whereas other line graphs of a similar type, namely $L(K_{n,n})$ for $n > 3$, are not? Following the analysis given in [22], we observe that for $n \leq 3$ these graphs are ultrahomogeneous – $L(K_{3,3})$ is the sporadic case, $L(K_{2,2})$ is a complete bipartite graph (i.e. one of the multipartite graphs in Theorem 1.4.2) and $L(K_{1,1})$ can be considered as the complete graph $K_1$. On the other hand, each graph $L(K_{n,n})$ for $n > 3$ contains two subgraphs, each isomorphic to $2 \cdot K_2$ (two disjoint edges), with the property that any isomorphism between them cannot be extended to an automorphism of whole graph. Figure 6.1 shows these subgraphs in $L(K_{4,4})$ – the first subgraph is represented by edges $v_1v_2$ and $u_1u_2$ and the second by edges $v_1v_2$ and $w_1w_2$.

This problem can be fixed by changing the structure under consideration – or equivalently by adding new relations to its signature [22]. For $K_{n,n}$, $n > 3$, we can add a relation of parallelism $P(v_1, v_2, v_3, v_4)$ which defines for every pair of disjoint edges with no additional edges between the four spanned vertices $v_1, v_2, v_3, v_4$ as being parallel if there is no vertex adjacent to each of vertices $v_1, v_2, v_3, v_4$. Figure 6.1 shows a parallel pair of edges $v_1v_2$ and $u_1u_2$, while edges $v_1v_2$ and $w_1w_2$ are not parallel due to the existence of a vertex $x$ adjacent to each of the endpoints. This extended definition of structure makes the whole class $L(K_{n,n})$ ultrahomogeneous. For more details see [22].

This process of making structures ultrahomogeneous, which is called homogenization, is also an important concept in the construction of universal structures [32, 73].

Throughout this chapter we denote graphs by boldface letters and their vertex set by the corresponding lightface letter, in a similar way to the notation used for general relational structures, in order to emphasize their interpretation as a relational structure admitting extensions by the adding of further relations. The only exceptions using standard notation as defined in Section 1.3 are particular graphs such as cycles or complete graphs and the Gaifman graph defined below.

## 6.2 Complexity of relational structures

We introduce our two notions of complexity: relational complexity and lift complexity. First, however, we review more formally the process of adding new types of relations into a structure.

Let $L$ be a signature containing relations $(R^i_{\mathbf{A}} : i \in I)$ with arities $(\delta_i : i \in I)$. Moreover let the signature $L'$ be an expansion of $L$. Then every structure $\mathbf{X} \in \mathrm{Rel}(L')$ may be viewed as a structure $\mathbf{A} = (A, (R^i_{\mathbf{A}} : i \in I)) \in \mathrm{Rel}(L)$ together with some additional relations for $i \in I' \setminus I$. To make this more explicit, these additional relations will be denoted by $X^i_{\mathbf{X}}, i \in I' \setminus I$. Thus a structure $\mathbf{X} \in \mathrm{Rel}(L')$ will be written as

$$\mathbf{X} = (A, (R^i_{\mathbf{A}} : i \in I), (X^i_{\mathbf{X}} : i \in I' \setminus I)),$$

and, abusing notation, more briefly as

$$\mathbf{X} = (\mathbf{A}, X^1_{\mathbf{X}}, X^2_{\mathbf{X}}, \ldots, X^N_{\mathbf{X}}).$$

We call $\mathbf{X}$ a *lift* of $\mathbf{A}$ and $\mathbf{A}$ is called a *shadow* of $\mathbf{X}$. In this sense the class $\mathrm{Rel}(L')$ is the class of all lifts of $\mathrm{Rel}(L)$. Conversely, $\mathrm{Rel}(L)$ is the class of all shadows of $\mathrm{Rel}(L')$. If all the extended relations are unary then the lift is called *monadic*. In the context of monadic lifts, the *color* of vertex $v$ representing a unary relation is the set $\{i : (v) \in X^i_{\mathbf{X}}\}$. Unless stated explicitly otherwise, we shall use letters $\mathbf{A}, \mathbf{B}, \mathbf{C}, \ldots$ for shadows (in $\mathrm{Rel}(L)$) and letters $\mathbf{X}, \mathbf{Y}, \mathbf{Z}$ for lifts (in $\mathrm{Rel}(L')$).

The *lift complexity*, $\mathrm{lc}(\mathbf{A})$, of a relational structure $\mathbf{A}$ is the least $k$ such that there exists a lift $\mathbf{X} = (\mathbf{A}, X^1_{\mathbf{X}}, X^2_{\mathbf{X}}, \ldots, X^N_{\mathbf{X}})$ of $\mathbf{A}$ that is ultrahomogeneous and all the relations $X^1_{\mathbf{X}}, X^2_{\mathbf{X}}, \ldots, X^N_{\mathbf{X}}$ have arities at most $k$.

Let $\mathbf{A}$ be a relational structure and let $\mathrm{Aut}(\mathbf{A})$ be the automorphism group of $\mathbf{A}$. A $k$-ary relation $R \subseteq A^k$ is an *invariant* of $\mathrm{Aut}(\mathbf{A})$ if $(\alpha(x_1), \ldots, \alpha(x_k)) \in R$ for all $\alpha \in \mathrm{Aut}(\mathbf{A})$ and all $(x_1, \ldots, x_k) \in R$. Let $\mathrm{Inv}_k(\mathbf{A})$ denote the set of all $k$-ary invariants of $\mathrm{Aut}(\mathbf{A})$ and let $\mathrm{Inv}(\mathbf{A}) = \bigcup_{k \geq 1} \mathrm{Inv}_k(\mathbf{A})$, $\mathrm{Inv}_{\leq k}(\mathbf{A}) = \bigcup_{1 \leq k' \leq k} \mathrm{Inv}_{k'}(\mathbf{A})$. It easily follows that the lift $(A, (R^i_{\mathbf{A}} : i \in I), \mathrm{Inv}(\mathbf{A}))$ (possibly with an infinite number of relations) is an ultrahomogeneous structure for every structure $\mathbf{A} = (A, (R^i_{\mathbf{A}} : i \in I))$. For a structure $\mathbf{A}$ the *relational complexity* $\mathrm{rc}(\mathbf{A})$ of $\mathbf{A}$ is the least $k$ such that $(A, (R^i_{\mathbf{A}} : i \in I), \mathrm{Inv}_{\leq k}(\mathbf{A}))$ is ultrahomogeneous, if such a $k$ exists. If no such $k$ exists, we say that the relational complexity of $\mathbf{A}$ is not finite and write $\mathrm{rc}(\mathbf{A}) = \infty$.

## 6.3 Basic properties

As a warm-up, we state a few basic observations about relational and lift complexities. It easily follows from the definition that the complement of an ultrahomogeneous structure is also ultrahomogeneous. The same holds for both of our notions of complexity.

**Proposition 6.3.1.** *Relational and lift complexity are each closed under complementation.*

Relational complexity is determined as for lift complexity except in the restrictions imposed on adding relations determined by the automorphism group of the structure. This immediately leads to the following simple fact.

**Proposition 6.3.2.** *For every structure* $\mathbf{A}$, $\mathrm{lc}(\mathbf{A}) \leq \mathrm{rc}(\mathbf{A})$.

Relational complexity is interesting even for finite structures, while lift complexity is trivially 1 for finite structures that are not ultrahomogeneous.

**Proposition 6.3.3.** *Let* $\mathbf{A}$ *be a finite relational structure. Then* $\mathrm{lc}(\mathbf{A}) \leq 1$ *and* $\mathrm{rc}(\mathbf{A}) \leq |A| - 1$.

*Proof.* For every finite $\mathbf{A}$, an ultrahomogeneous lift can be created by adding a unique unary relation to each vertex. This gives $\mathrm{lc}(\mathbf{A}) \leq 1$.

The unary relations may not be invariant. It can however be easily seen that by adding all invariant relations (i.e. those having arity at most $|A|$) one always obtains an ultrahomogeneous structure. A relational structure of arity $|A|$ does not however contribute to the homogenization of the structure, giving the bound of $|A| - 1$ on the relational complexity of a finite structure. □

The following observation allows us to restrict our attention to connected structures.

**Proposition 6.3.4.** *Let* $k \geq 2$ *be finite and let* $\mathbf{A}$ *be a non-ultrahomogeneous relational structure with connected components* $\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_k$. *Then*

1. $\mathrm{lc}(\mathbf{A}) = \max\{1, \mathrm{lc}(\mathbf{A}_1), \mathrm{lc}(\mathbf{A}_2), \ldots \mathrm{lc}(\mathbf{A}_k)\}$;
2. *if there is a pair of two mutually isomorphic structures* $\mathbf{A}_i$ *and* $\mathbf{A}_j$, $i \neq j$, *such that* $rc(\mathbf{A}_i + \mathbf{A}_j) > 1$, *then*

$$\mathrm{rc}(\mathbf{A}) = \max\{2, \mathrm{rc}(\mathbf{A}_1), \mathrm{rc}(\mathbf{A}_2), \ldots \mathrm{rc}(\mathbf{A}_k)\},$$

*otherwise*

$$\mathrm{rc}(\mathbf{A}) = \max\{1, \mathrm{rc}(\mathbf{A}_1), \mathrm{rc}(\mathbf{A}_2), \ldots \mathrm{rc}(\mathbf{A}_k)\}.$$

(Here $\mathbf{A} + \mathbf{B}$ denotes the disjoint union of $\mathbf{A}$ and $\mathbf{B}$.)

*Proof.* The ultrahomogeneous lift of the structure $\mathbf{A}$ can always be created as a disjoint union of ultrahomogeneous lifts of $\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_k$ with $k$ additional unary relations distinguishing individual components. With this construction the lift complexity is increased to at least 1. This finishes the proof of (1).

To show (2) we need only add invariant unary and binary relations depending on the isomorphism types of components. First we add unary relations classifying vertices by the isomorphism type of the connected component to which they belong. If there is no pair of mutually isomorphic components $\mathbf{A}_i, \mathbf{A}_j$ such that $\mathrm{rc}(\mathbf{A}_i + \mathbf{A}_j) > 1$, then the lift is extended by all necessary relations that make each component ultrahomogeneous, which creates an ultrahomogeneous lift.

For every $\mathbf{A}_i$ along with one or more components $\mathbf{A}_{j_1}$, $\mathbf{A}_{j_2}$, $\ldots \mathbf{A}_{j_n}$ that are all distinct and isomorphic to $\mathbf{A}_i$, we add two extra binary relations. First a relation used to join all pairs of vertices within the same component, and second a relation joining all pairs of vertices belonging to two different components.

These additional relations prevent partial isomorphisms exchanging vertices in between individual connected components leading to an ultrahomogeneous structure after applying all remaining relations in a similar way to above. □

### 6.3.1 Graphs of complexity 1

Our notion of relational complexity of structures is derived from the notion of relational complexity of groups in [22]. For every structure $\mathbf{A}$, the relational complexity of $\mathrm{Aut}(\mathbf{A})$ (in the sense of [22]) corresponds to $\mathrm{rc}(\mathbf{A})$ with the exception of $\mathrm{rc}(\mathbf{A})$ being smaller than the maximum arity of a relation in $\mathbf{A}$. Our notion of relational complexity ignore arities of relations already present in $\mathbf{A}$. For instance, the relational complexity of the cyclic group on $n \geq 4$ elements in the sense of [22] is 2, while $\mathrm{rc}(C_4)$ and $\mathrm{rc}(C_5)$ is 0. The main motivation for our definition is to get finer information on structures with small complexity. We explore these small complexity classes now. To simplify our presentation we shall restrict our attention to graphs. (Many of our observations generalize to arbitrary relational structures.)

The class of graphs of relational or lift complexity 0 has been well studied: these are the ultrahomogeneous graphs. We now consider the class of graphs of relational complexity 1 and the class of graphs of lift complexity 1. Both these clases are closely related to the established notion of an $n$-graph (see [133] for a recent review of the topic).

**Definition 6.3.1.** For positive integers $n$ and $c$, a *$c$-colored $n$-graph* is a graph on $n$ pairwise disjoint sets of vertices $V_1, V_2, \ldots, V_n$ (called *parts*), each of which is an ordinary countable graph, with $c$ edge-types between pairs of parts (cross-edges).

Isomorphisms of $n$-graphs do not permit exchanging vertices within an individual part [133]. We shall consider 2-colored $n$-graphs as corresponding to graphs where one of the types of cross-edges is an edge and the other type is a non-edge.

The unary relations forming an ultrahomogeneous lift (equivalently seen as a vertex coloring) partitions the vertex set of a graph into a finite number of classes. Analogously to $n$-graphs we call the classes of this partition *parts*.

The following observation describes the structure of graphs of complexity 1.

**Proposition 6.3.5.** *Let $\mathbf{G}$ be a graph with $\mathrm{rc}(\mathbf{G}) = 1$ or $\mathrm{lc}(\mathbf{G}) = 1$ and let $V_1, V_2, \ldots, V_k$ be its parts. Then the following holds.*

1. *The subgraph induced by each part is an ultrahomogeneous graph.*
2. *$\mathbf{G}$ corresponds to an ultrahomogeneous 2-colored $k$-graph with parts $V_1, V_2, \ldots, V_k$.*
3. *The subgraph induced by each pair of parts corresponds to an ultrahomogeneous 2-colored 2-graph.*

*Proof.* The automorphisms of $\mathbf{G}$ consist of arbitrary combinations of automorphisms of subgraphs induced by the individual parts. Consequently, each of the subgraphs must already be ultrahomogeneous, giving (1).

(2) and (3) follow directly from ultrahomogeneity. □

The structural condition given by Proposition 6.3.5 and the classification of ultrahomogeneous graphs allows the construction of a number of interesting examples. Generally, $c$-colored $n$-graphs are only partially classified. The classification is complete for the class of *$n$-edge-colored bipartite* graphs, which is a special subclass of $n$-colored 2-graphs where every part is an independent set.

**Theorem 6.3.6** ([78]). *If* $\mathbf{G}$ *is a countable ultrahomogeneous* $n$*-edge-colored bipartite graph such that* $n$ *is finite and all of the* $n$ *types of edges are used in* $\mathbf{G}$*, then one of the following holds:*

- $n = 1$ *and all edges are the same color;*
- $n = 2$ *and edges of one color form a perfect matching, and edges of the other color are its complement;*
- $n \geq 2$ *and* $\mathbf{G}$ *is a generic bipartite* $n$*-edge-colored graph.*

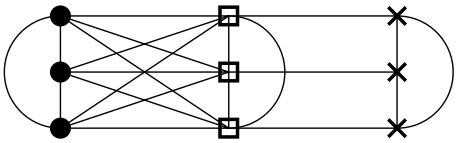


Figure 6.2: Graph of relational complexity 1 consisting of 3 isomorphic ultrahomogeneous graphs ($K_3$). The parts are depicted by different vertex markings.

In the case of graphs ($n \leq 2$), we thus have only three types of bipartite ultrahomogeneous graphs (up to complementation): complete or empty bipartite graphs, matchings, and the generic bipartite graph. In Figure 6.2 we show a graph of relational complexity 1 that consists of 3 isomorphic ultrahomogeneous subgraphs.

Observe that not every graph built from finitely many isomorphic copies of an ultrahomogeneous graph has relational complexity 1. In the example given it is the use of different bipartite graphs to connect each pair of subgraphs that ensures that each of the subgraphs forms an independent cluster. Proposition 6.3.5 can however be reversed for lift complexity:

**Proposition 6.3.7.** *Every infinite graph* $\mathbf{G}$ *with* $n$ *parts (*$n$ *finite) such that* $\mathrm{lc}(\mathbf{G}) = 1$ *corresponds to an ultrahomogeneous* 2*-colored* $n$*-graph.*

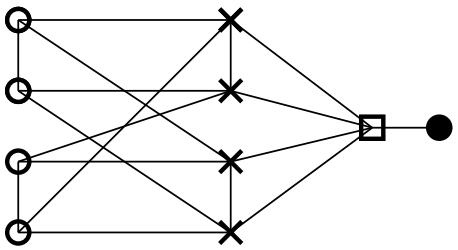


Figure 6.3: Graph of relational complexity 1 using a sporadic 2-colored 2-graph.

The classification of ultrahomogeneous 2-colored $n$-graphs is still an open problem – see [133] for partial results. The graph depicted in Figure 6.3 has relational complexity 1 and is constructed from 4 ultrahomogeneous subgraphs ($K_1$, $K_1$, $K_2+K_2$ and $K_2+K_2$) using a sporadic example of an ultrahomogeneous 2-colored 2-graph joining $K_2 + K_2$ and $K_2 + K_2$ (c.f. [133]).

A special case of graphs of lift complexity 1 can be constructed from a given finite graph $\mathbf{G}$ by replacing every vertex by an ultrahomogeneous graph and

every edge by an ultrahomogeneous 2-colored 2-graph. Such a construction is considered in [73], where the homomorphism dual $\mathbf{D}$ of a graph $\mathbf{T}$ is turned into a countable graph $\mathbf{U}$ that is embedding-universal for the class of all graphs not containing a homomorphic image of $\mathbf{T}$. Here vertices of $\mathbf{D}$ are replaced by infinite discrete graphs and edges of $\mathbf{D}$ by random bipartite graphs. It is a well-known result [117] that homomorphism duals exists only when $\mathbf{T}$ is a (relational) tree. For the case of universal graphs with lift complexity 1, the construction can be extended also to structures obtained from trees by replacing every edge by an arbitrary irreducible structure. When the same construction is applied to any rigid graph $\mathbf{D}$ (and this is the case for cores of graph duals), the expanded graph also has relational complexity 1. We shall further study the complexity of infinite graphs in greater detail in Section 6.5.

### 6.3.2 Graphs of complexity 2

Given the difficulties of characterizing graphs even of complexity 1, it is not reasonable to expect a simple characterization of graphs of complexity 2. We can however produce several interesting classes of such graphs.

Given a graph $\mathbf{G}$, the *graph metric* is a function measuring the path distance between pairs of vertices. We call a graph *metrically ultrahomogeneous* if and only if it is ultrahomogeneous as a metric space with its graph metric. It is not difficult to see that relations measuring distance between vertices (i.e. relations connecting vertices of distance $n$) are all invariant binary relations. We immediately have the following observation.

**Proposition 6.3.8.** *All metrically ultrahomogeneous graphs have relational and lift complexity at most 2.*

Metrically ultrahomogeneous graphs were studied by Cherlin, leading to a partial catalog of them [23]. These include special bipartite graphs, tree-like graphs, antipodal graphs, and a number of other examples.

One example of a connected graph of relational complexity 2 that is not metrically ultrahomogeneous can be constructed with the help of Proposition 6.3.4. Take the complement of the graph created as $C_5 + C_5$. This graph has relational complexity 2, but it is not metrically ultrahomogeneous. Consider a function mapping an edge within the complement of $C_5$ to an edge joining the two subgraphs. Additional examples can be easily produced with the help of the following two propositions.

**Proposition 6.3.9.** *Finite (graph) trees have relational complexity at most* 2.

*Proof.* We prove this by induction on the diameter of a given tree $\mathbf{T}$. Again we shall consider the unary relations to be vertex colorings. To carry through the induction, we prove a stronger result: all finite vertex-colored trees have relational complexity at most 2.

The claim trivially holds for a colored tree consisting of a single vertex or edge.

Now assume that the claim holds for trees of diameter up to $k$. Fix a colored tree $\mathbf{T}$ of diameter $k+2$ and consider a tree $\mathbf{T}'$ constructed from $\mathbf{T}$ by

1. removing all leaves; and

2. changing the color of every new leaf vertex $v$ to a unique representation of the isomorphism type of the rooted tree induced by $\mathbf{T}$ on $v$ and its descendants.

$\mathbf{T}'$ is a vertex-colored tree with diameter $k$. By induction hypothesis $\mathrm{rc}(\mathbf{T}') \leq 2$ and thus there exists an ultrahomogeneous lift $\mathbf{X}'$ of $\mathbf{T}'$ adding only invariant unary and binary relations. To make our presentation easier, we also consider the colors introduced in the construction of $\mathbf{T}'$ to be extended unary relations of $\mathbf{X}'$ and edges of $\mathbf{T}'$ to be mirrored in an extended binary relation of $\mathbf{X}'$.

We extend $\mathbf{X}'$ to an ultrahomogeneous lift $\mathbf{X}$ of $\mathbf{T}$. Once again we extend our language of $\mathbf{X}$, in the following way:

1. for every unary relation $X^i_{\mathbf{X}'}$ we add a new unary relation $X^{u(i)}_{\mathbf{X}}$ and a new binary relation $X^{b(i)}_{\mathbf{X}}$; and

2. for every binary relation $X^i_{\mathbf{X}'}$ we add a new binary relation $X^{b(i)}_{\mathbf{X}}$.

$\mathbf{X}$ is a lift of $\mathbf{T}$ such that $X^i_{\mathbf{X}} = X^i_{\mathbf{X}'}$ for all relations used by $\mathbf{X}'$. We use the newly introduced relations in the following way.

1. for every leaf $v \in T$ with father $v'$ such that $(v') \in X^i_{\mathbf{X}'}$ we also put $(v) \in X^{u(i)}_{\mathbf{X}}$;

2. for every pair of distinct leaves $v_1, v_2 \in T$ sharing father $v'$ such that $(v') \in X^i_{\mathbf{X}'}$ we also put $(v_1, v_2) \in X^{b(i)}_{\mathbf{X}}$ and $(v_2, v_1) \in X^{b(i)}_{\mathbf{X}}$;

3. for every pair of distinct leaves $v_1, v_2 \in T$ with distinct fathers $v'_1, v'_2$ (respectively) such that $(v'_1, v'_2) \in X^i_{\mathbf{X}'}$ we also put $(v_1, v_2) \in X^{b(i)}_{\mathbf{X}}$.

The ultrahomogeneity of $\mathbf{X}'$ follows from the fact that automorphisms of $\mathbf{T}$ must map leaves to leaves and non-leaves to non-leaves. The automorphism group of $\mathbf{T}$ acting on non-leaf vertices of $\mathbf{T}$ is precisely the automorphism group of $\mathbf{X}'$. Finally, every automorphism can map a leaf vertex $v$ to a leaf vertex $v'$ if and only if they have same color and it can map the father of vertex $v$ to the father of vertex $v'$.

□

Because graphs with relational complexity 2 are closed under complementation, Proposition 6.3.4 gives an iterative way to construct non-trivial examples of such graphs. Consider the following special case. A *cograph* is a graph not containing an induced path on 4 vertices. It is well-known that all cographs can be generated from the single-vertex graph $K_1$ by complementation and disjoint union. We immediately obtain:

**Proposition 6.3.10.** *Finite cographs have relational complexity at most 2.*

## 6.4 Finite graphs with large complexity

It is not difficult to construct examples of finite graphs with large relational complexity, showing that the relational complexity of graphs is unbounded.

**Example 6.4.1.** Consider a permutation group $\Gamma$ acting on $n$ elements that is $k$-transitive but not $(k+1)$-transitive. (If $\Gamma$ is the alternating group on $n$ elements, then $k = n-2$.) Now construct a graph $\mathbf{G}_\Gamma$ with the following vertices:

1. $n$ control vertices $v_1, v_2, \dots v_n$;
2. for every permutation $p \in \Gamma$, $n+1$ additional vertices $v_1^p, v_2^p, \dots, v_{n+1}^p$;

and the following edges:

1. $\{v_a^p, v_{a+1}^p\}$ for every $p \in \Gamma$ and $1 \leq a \leq n$;
2. $\{v_i^p, v_j\}$ if and only if the permutation $p \in \Gamma$ sends the $i$th element to the $j$th element.

The graph $\mathbf{G}_\Gamma$ thus consists of control vertices and paths representing individual permutations connected to the control vertices by pairings. By construction, an automorphism of $\mathbf{G}_\Gamma$ sends control vertices to control vertices (these are the only vertices of large degree). Similarly, an automorphism must send a path $v_1^p, v_2^p, \dots, v_{n+1}^p$ corresponding to some permutation $p \in \Gamma$ to another path $v_1^{p'}, v_2^{p'}, \dots, v_{n+1}^{p'}$ corresponding to another permutation $p' \in \Gamma$. It easily follows that the automorphism group of $\mathbf{G}_\Gamma$ acting on the control vertices is precisely $\Gamma$. By transitivity of $\Gamma$ it is necessary to use at least $(k+1)$-ary relations in order to homogenize it, which gives $\mathrm{rc}(\mathbf{G}_\Gamma) > k$. We remark that this construction in fact works for all groups with large relational complexity in the sense of [22].

Identifying less artificial families of graphs with large relational complexity is however challenging. We outline two examples given by [22].

1. The *Johnson graph* $J_{n,k}$ is a graph whose vertices are all the $k$-subsets of a fixed $n$ element set. Two vertices are adjacent when the two corresponding sets meet in exactly $k-1$ elements. In [29] the bound $\mathrm{rc}(J_{n,k}) \leq 2\lceil \log_2 k \rceil$ is given on the relational complexity of Johnson graphs. Equality is achieved when $n \geq 2\log_2 k + 2$.
2. The *Kneser graph* $KG_{n,k}$ is the graph whose vertices are the $k$-subsets of a fixed set of $n$ elements, and where two vertices are adjacent if and only if the two corresponding sets are disjoint. For $n \geq 2k$ the relational complexity is given by $\mathrm{rc}(KG_{n,k}) = 2\lceil \log_2 k \rceil$.

   In particular, the relational complexity of the Petersen graph is 3 and thus the problem shown in Figure 1.2 is in fact the only obstacle to its ultrahomogeneity.

It seems that the upper bound on relational complexity given by Proposition 6.3.3 is far from the reality, since the relational complexity of all the examples seen so far grows sublogarithmically in their size. It is reasonable to ask for extremal examples of graphs with high relational complexity or for better bounds.

**Problem 6.4.1.** *Estimate* $f(n) = \max_{|G|=n} \mathrm{rc}(G)$.

## 6.5 Complexity of infinite structures

We have primarily discussed the relational complexity of finite structures. Now we turn our attention to the infinite case, as has been proposed in [59]. Fraïssé's Theorem 1.2.8 can be seen as a "zero instance" of the problems considered in this section. In this theorem the amalgamation property can be seen as the critical property that the age $\mathcal{K}$ must have in order that there exists structure $\mathbf{U}$ with $\mathrm{Age}(\mathbf{U}) = \mathcal{K}$ satisfying $\mathrm{rc}(\mathbf{A}) = 0$. Inspired by this approach, we could, for a fixed age $\mathcal{K}$, ask for bounds on relational complexity for structures $\mathbf{U}$ having $\mathrm{Age}(\mathbf{U}) = \mathcal{K}$. Alternatively, fixing $n$, we seek structural properties of the age $\mathcal{K}$ that imply the existence of a structure $\mathbf{U}$ with $\mathrm{Age}(\mathbf{U}) = \mathcal{K}$ and $\mathrm{rc}(\mathbf{U}) = n$.

Relational complexity is not interesting for rigid structures (with trivial automorphism group), where it is always 1. Such a structure exists for almost every age. We thus restrict our attention to $\omega$-categorical structures. Recall that using Theorem 1.5.1, proved independently by Engeler, Ryll-Nardzewski and Svenonius, we can characterize $\omega$-categorical structures as structures whose automorphism group has only finitely many orbits on $n$-tuples, for every $n$, and thus there are also only finitely many invariant relations of arity $n$.

For this section several versions of amalgamation are needed and it is thus useful to give them standard names. Suppose we have an amalgamation of structures $\mathbf{A}$, $\mathbf{B}_1$ and $\mathbf{B}_2$, as given in Definition 1.2.3. If the embeddings $f_1$ and $f_2$ are identity mappings we call this amalgamation an *amalgamation of* $\mathbf{B}_1$ *and* $\mathbf{B}_2$ *over* $\mathbf{A}$. An amalgamation $(\mathbf{B}_1, \mathbf{B}_2, \mathbf{A}, f_1, f_2)$ is called *strong* if

$$g_1(B_1) \cap g_2(B_2) = g_1(f_1(A)) = g_2(f_2(A)).$$

Moreover we call a strong amalgamation *free* if there are no relations of $\mathbf{C}$ spanning vertices of both $g_1(B_1)$ and $g_2(B_2)$ that are not images of some relations of structure $\mathbf{B}_1$ or $\mathbf{B}_2$ via the embedding $g_1$ or $g_2$, respectively.

Using Theorem 1.5.2 we can see that the countable $\omega$-categorical structure $\mathbf{U}$ is universal for the class of all countable structures younger than $\mathbf{A}$ and thus it contains as an induced substructure every countable structure $\mathbf{A}$, $\mathrm{Age}(\mathbf{A}) \subseteq \mathrm{Age}(\mathbf{U})$. However there is no 1-1 correspondence between $\omega$-categorical structures and their ages. We will demonstrate this for the class of bipartite graphs.

Consider the class $\mathcal{K}$ of lifts of finite bipartite graphs with one part distinguished by an extended unary relation. $\mathcal{K}$ is an amalgamation class and thus there is an (up to isomorphism unique) infinite ultrahomogeneous lift $\mathbf{X}, \mathrm{Age}(\mathbf{X}) = \mathcal{K}$. Now consider a bipartite graph $\mathbf{B}_2$ that is a shadow of $\mathbf{X}$. The graph $\mathbf{B}_2$ is not ultrahomogeneous, but is however a universal bipartite graph (because $\mathbf{X}$ is universal for bipartite graphs with one part distinguished). Each vertex of $\mathbf{B}_2$ has infinite degree and $\mathbf{B}_2$ is connected. The age of $\mathbf{B}_2$ is the class of all finite bipartite graphs. Every countable bipartite graph is an induced subgraph of $\mathbf{B}_2$ (it is a universal bipartite graph). There are no non-trivial invariant unary relations because $\mathbf{B}_2$ is vertex transitive. It is possible to turn $\mathbf{B}_2$ into an ultrahomogeneous lift by using two invariant binary relations. The first relation joins every pair of vertices belonging to the same part. The second relation joins every pair of vertices belonging to different parts. Consequently $\mathrm{rc}(\mathbf{B}_2) = 2$ while $\mathrm{lc}(\mathbf{B}_2) = 1$.

Denote by $L$ the set of vertices of $\mathbf{X}$ belonging to the part distinguished by the unary relation. Now consider the graph $\mathbf{B}_1$ created from $\mathbf{X}$ by joining every

$v \in L$ to a newly added vertex of degree 1. An automorphism of $\mathbf{B}_1$ necessarily maps vertices of degree 1 to vertices of degree 1 and thus cannot swap the parts. It is thus possible to turn $\mathbf{B}_1$ into an ultrahomogeneous structure with two unary relations and $\mathrm{rc}(\mathbf{B}_1) = \mathrm{lc}(\mathbf{B}_1) = 1$. The age is however unchanged.

Finally it is possible to construct, for given $n > 2$, a connected bipartite universal graph $\mathbf{B}_n$, $\mathrm{rc}(\mathbf{B}_n) = n$. Take any connected bipartite graph $\mathbf{A}_n$, $\mathrm{rc}(\mathbf{A}_n) = n$. (Such a graph can be constructed by the techniques of Section 6.4; non-bipartite graphs can be turned into bipartite ones by subdividing every edge by a vertex). Moreover we have already constructed universal random bipartite graph $\mathbf{B}_2$. Construct $\mathbf{B}_n$ as a disjoint union of $\mathbf{B}_2$ and $\mathbf{A}_n$ with one vertex unified.

It follows that an $\omega$-categorical relational structure whose age consists of all finite bipartite graphs can have relational complexity anywhere between 1 and infinity. This is not a sporadic example and we thus need to add extra restrictions on the structures under consideration. Among all $\omega$-categorical structures with a given age we can turn our attention to the "most ultrahomogeneous-like" in the following sense. A structure $\mathbf{A}$ with $\mathrm{Age}(\mathbf{A}) = \mathcal{K}$ is *existentially complete* if for every structure $\mathbf{B}$, such that $\mathrm{Age}(\mathbf{B}) = \mathcal{K}$ and the identity mapping (on $A$) is an embedding $\mathbf{A} \to \mathbf{B}$, every existential statement $\psi$ which is defined in $\mathbf{A}$ and true in $\mathbf{B}$ is also true in $\mathbf{A}$. By [25], for every age $\mathcal{K}$ defined by forbidden monomorphisms with an $\omega$-categorical universal structure there is also up to isomorphism a unique $\omega$-categorical, existentially complete and $\omega$-saturated universal structure, where $\omega$-saturated structure is understood as defined in [67]. This in fact holds more generally. In such cases the *canonical universal structure* of a given age $\mathcal{K}$ is the unique $\omega$-categorical, existentially complete, and $\omega$-saturated structure $\mathbf{U}$ such that $\mathrm{Age}(\mathbf{U}) = \mathcal{K}$.

Given an age $\mathcal{K}$ we can thus ask:

I. What is the minimal relational complexity of an $\omega$-categorical structure $\mathbf{U}$ such that $\mathrm{Age}(\mathbf{U}) = \mathcal{K}$?

II. What is the relational complexity of the canonical universal structure of age $\mathcal{K}$?

We consider universal structures for the class $\mathrm{Forb_h}(\mathcal{F})$, where $\mathcal{F}$ is a family of connected structures. Recall that $\mathrm{Forb_h}(\mathcal{F})$ denotes the class of all structures $\mathbf{A}$ for which there is no homomorphism $\mathbf{F} \to \mathbf{A}$, $\mathbf{F} \in \mathcal{F}$. The classes $\mathrm{Forb_h}(\mathcal{F})$ are among the most natural ones where the existence of a universal structure is guaranteed for every finite $\mathcal{F}$, see [25]. For such $\mathcal{F}$ we can fully answer the questions above.

For a structure $\mathbf{A} = (A, (R^i_{\mathbf{A}}, i \in I))$, the *Gaifman graph* (in combinatorics often called the *2-section*) is the graph $G_{\mathbf{A}}$ with vertices $A$ and edges all pairs that are contained in a tuple of a relation of $\mathbf{A}$, i.e., $G = (A, E)$, where $x, y \in E$ if and only if $x \neq y$ and there exists a tuple $\vec{v} \in R^i_{\mathbf{A}}$, $i \in I$, such that $x, y \in \vec{v}$.

For a structure $\mathbf{A}$ and a subset of its vertices $B \subseteq A$, the *neighborhood* of $B$ is the set of all vertices of $A \setminus B$ connected in $G_{\mathbf{A}}$ by an edge to a vertex of $B$. We denote by $G_{\mathbf{A}} \setminus B$ the graph created from $G_{\mathbf{A}}$ by removing the vertices in $B$. A structure $\mathbf{A}$ is called *connected* if its Gaifman graph $G_{\mathbf{A}}$ is connected.

A *g-cut* in $\mathbf{A}$ is a subset $C$ of $A$ which disconnects the Gaifman graph $G_{\mathbf{A}}$ when removed. A g-cut $C$ is *minimal g-separating* in $\mathbf{A}$ if there exists structures

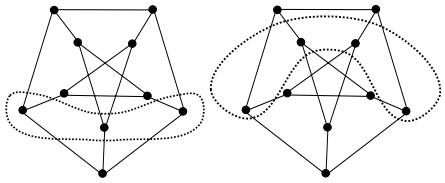


Figure 6.4: Two minimal g-separating g-cuts of the Petersen graph.

$\mathbf{A}_1 \neq \mathbf{A}_2$ induced by $\mathbf{A}$ on two connected components of $G_{\mathbf{A}} \setminus C$ such that $C$ is the intersection of the neighborhood of $A_1$ and the neighborhood of $A_2$ in $\mathbf{A}$. All the minimal separating cuts of the Petersen graph (up to isomorphism) are depicted in Figure 6.4.

A family of structures is called *minimal* if and only if all the structures in $\mathcal{F}$ are cores and there is no homomorphism between any two structures in $\mathcal{F}$. The following theorem provides bounds on the relational complexity of universal structures defined by a forbidden homomorphism class for various classes $\mathcal{F}$.

**Theorem 6.5.1.** *Let $\mathcal{F}$ be a finite minimal family of finite connected relational structures and $\mathbf{U}$ an $\omega$-categorical universal structure for $\mathrm{Forb_h}(\mathcal{F})$. Denote by $n$ the size of the largest minimal g-separating g-cut in $\mathcal{F}$. Then (a) $\mathrm{rc}(\mathbf{U}) \geq n$; (b) if $\mathbf{U}$ is the canonical universal structure for $\mathrm{Forb_h}(\mathcal{F})$ then $\mathrm{rc}(\mathbf{U}) = n$.*

In the rest of the chapter we establish the upper bounds and lower bounds given in Theorem 6.5.1. In fact we prove more general statements.

### 6.5.1 Upper bounds on relational complexity

It appears that relational complexity is closely related to the homogenization method of constructing universal structures as used in [32]. The main result of [32] is in fact a variant of Fraïssé's Theorem with amalgamation reduced to what is called local failure of amalgamation.

An *amalgamation failure* of a given age $\mathcal{K}$ is a triple $(\mathbf{A}, \mathbf{B}, \mathbf{C})$ such that $\mathbf{A}, \mathbf{B}, \mathbf{C} \in \mathcal{K}$, the identity mapping (on $C$) is an embedding $\mathbf{C} \to \mathbf{A}$ and $\mathbf{C} \to \mathbf{B}$, and there is no amalgamation of $\mathbf{A}$ and $\mathbf{B}$ over $\mathbf{C}$ in $\mathcal{K}$. (i.e., $(\mathbf{A}, \mathbf{B}, \mathbf{C})$ shows that $\mathcal{K}$ does not have the amalgamation property). An amalgamation failure is *minimal* if there is no another amalgamation failure $(\mathbf{A}', \mathbf{B}', \mathbf{C}')$ such that identity mappings are embeddings $\mathbf{A}' \to \mathbf{A}$, $\mathbf{B}' \to \mathbf{B}$ and $\mathbf{C}' \to \mathbf{C}$.

**Theorem 6.5.2.** *Let $\mathbf{U}$ be the canonical universal structure for age $\mathcal{K}$ and $S$ the set of isomorphism types of minimal amalgamation failures of $\mathbf{U}$. If $S$ is finite then $\mathrm{rc}(\mathbf{U})$ and $\mathrm{lc}(\mathbf{U})$ is bounded from above by the largest size of $\mathbf{C}$ such that $(\mathbf{A}, \mathbf{B}, \mathbf{C}) \in S$.*

*Proof.* Given an age $\mathcal{K}$ and set $S$, [32] provides a construction of an ultrahomogeneous lift $\mathbf{X}$ such that its shadow $\mathbf{U}$ is universal for the class of structures of age $\mathcal{K}$. Moreover this lift is constructed using relations invariant under the automorphism group of $\mathbf{U}$ (Lemma 2.7 in [32]) and their arities correspond to the sizes of $\mathbf{C}$ such that $(\mathbf{A}, \mathbf{B}, \mathbf{C}) \in S$. $\mathbf{U}$ is $\omega$-categorical because $\mathbf{U}$ is a shadow of $\mathbf{X}$. Existential completeness and $\omega$-saturation follow directly from the construction. □

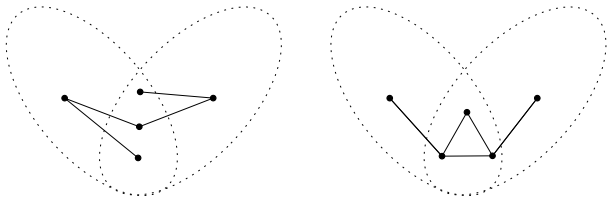


Figure 6.5: Amalgamation failures of the class of graphs not containing an induced path of length 3.

**Example 6.5.1.** Determining the minimal set of amalgamation failures of a given age can be difficult. In [32] there is only one non-trivial example given, namely that of cographs (graphs without an induced path on 4 vertices). The minimal amalgamation failures are depicted in Figure 6.5. This class has only two minimal failures, both containing 3 vertices in $\mathbf{C}$. The resulting lift uses a relation of arity 3. This is in contrast with the fact that all finite cographs have complexity at most 2 (Proposition 6.3.10).

In the special case of $\mathcal{K} = \mathrm{Age}(\mathrm{Forb_h}(\mathcal{F}))$ one can however prove a stronger result. Details of the following constructions will appear in [58]. Some of the construction used in this proof has also appeared in [72] as part of an informal cooperation with the respective authors. Before going any further several notions have to be introduced. First, it is useful to introduce substructures with a specific part "participating" in a g-cut. For any set of elements $A$ denote by $\vec{A}$ a tuple of corresponding vertices. For any relational structure $\mathbf{A}$ we can define a *rooted structure* $\mathcal{A}$ as a pair $(\mathbf{A}, \vec{R})$, where $\vec{R}$ is a tuple of distinct vertices of $\mathbf{A}$ called a *root*. Since roots are going to be used for amalgamations, which operate with isomorphisms, we call the size of this root the *width* of the rooted structure. This also plays its role in isomorphism between rooted structures. We say that rooted structures $\mathcal{A}_1 = (\mathbf{A}_1, \vec{R}_1)$ and $\mathcal{A}_2 = (\mathbf{A}_2, \vec{R}_2)$ are isomorphic if there is an isomorphism $f : A_1 \to A_2$ between $\mathcal{A}_1$ and $\mathcal{A}_2$ such that $f|_{\vec{R}_1}$ is a monotone bijection taking $\vec{R}_1$ to $\vec{R}_2$.

Let $\mathbf{A}$ be a connected relational structure and $R$ a minimal g-separating g-cut forming a component $C$ of the Gaifman graph $G_{\mathbf{A}}$. A *piece* of relational structure $\mathbf{A}$ is a rooted structure $(\mathbf{P}, \vec{R})$, where the tuple $\vec{R}$ consists of the vertices of the g-cut $R$ in a fixed linear order and $\mathbf{P}$ is the structure induced by $\mathbf{A}$ on $C \cup R$.

To construct universal structures special types of lift have to be introduced. For a family of relational structures $\mathcal{F}$ we fix an enumeration of pieces $\mathcal{P}_i = (\mathbf{P}_i, \vec{R}_i), i \in I'$. For a relational structure $\mathbf{A}$ define the lift $\mathbf{X} = (\mathbf{A}, (X^i_{\mathbf{X}}, i \in I'))$. This lift is then called an *$\mathcal{F}$-lift* if the arities of the relations $X^i_{\mathbf{X}}, i \in I'$ correspond to the sizes of the roots in the given enumeration, i.e. $|\vec{R}_i|$. The $\mathcal{F}$-lift $\mathbf{X}$ is then called the *canonical lift*, denoted by $\mathbf{X} = L(\mathbf{A})$, when for any tuple of size $k$ it is the case that $(x_1, x_2, \ldots, x_k) \in X^i_{\mathbf{X}}$ if and only if there exists a homomorphism $f$ from $\mathbf{P}_i$ to $\mathbf{A}$ such that $f(\vec{R}_i) = (x_1, x_2, \ldots, x_k)$. In other words, the lift relations glue only those tuples that are contained in the corresponding piece.

Let $\mathbf{A}$ be a relational structure and $B \subseteq A$. We call the canonical lift $L(\mathbf{A})$ *complete on* $B$ if for every $\mathbf{C} \in \mathrm{Forb_h}(\mathcal{F})$ such that $\mathbf{A} \subseteq \mathbf{C}$ the lift induced by $L(\mathbf{A})$ on $B$ is the same as the lift induced by $L(\mathbf{C})$ on $B$. Finally, the lift $\mathbf{X}$ is called *complete* if there exists $\mathbf{C} \in \mathrm{Forb_h}(\mathcal{F})$ such that $\mathbf{X}$ is induced by $L(\mathbf{C})$ on

$X$ and the canonical lift $L(\mathbf{C})$ is complete on $X$.

Let $\mathcal{L}_\mathcal{F}$ denote the set of all complete lifts of $\mathrm{Forb_h}(\mathcal{F})$. For a complete lift $\mathbf{X} \in \mathcal{L}_\mathcal{F}$ we define a *witness* $W(\mathbf{X})$ of the fact that $\mathbf{X}$ belongs to $\mathcal{L}_\mathcal{F}$ as a relational structure $\mathbf{A} \in \mathrm{Forb_h}(\mathcal{F})$ such that $\mathbf{X}$ is induced by $L(\mathbf{A})$ on $X$ and $L(\mathbf{A})$ is complete on $X$. The inclusion of completeness in the definition of witness ensures that constructing a universal structure is existentially complete. Moreover free amalgamation preserves the property of being a witness, as stated in the following lemma, whose proof we omit.

**Lemma 6.5.3** ([72])**.** *Let* $\mathbf{A}$ *and* $\mathbf{B}$ *be both witnesses of* $\mathbf{X}$*. Then the free amalgam of* $\mathbf{A}$ *and* $\mathbf{B}$ *over* $\mathbf{X}$ *is also a witness.*

We now have all the machinery needed in order to state the main theorem bounding both complexities. A complete proof with all necessary details will appear in [58] and as mentioned above some of the construction used in this proof has also appeared in [72] as part of an informal cooperation with the authors of the cited paper.

**Theorem 6.5.4.** *Let* $\mathcal{F}$ *be a (finite or infinite) family of connected structures such that there exists a canonical universal structure* $\mathbf{U}$ *for* $\mathrm{Age}(\mathrm{Forb_h}(\mathcal{F}))$*. Then* $\mathrm{rc}(\mathbf{U})$ *and* $\mathrm{lc}(\mathbf{U})$ *are each bounded from above by the size of the largest minimal g-separating g-cut in* $\mathcal{F}$*.*

*Proof.* Let $\mathcal{K} = \mathrm{Age}(\mathrm{Forb_h}(\mathcal{F}))$. By assumption the canonical universal structure $\mathbf{U}$ exists. Define $\mathcal{K}_\mathcal{F}$ to be the class of all complete lifts of $\mathrm{Age}(\mathrm{Forb_h}(\mathcal{F}))$. By definition of $\mathcal{K}$, the class $\mathcal{K}_\mathcal{F}$ obviously satisfies both properties **HP** and **JEP**. It remains to prove it has the amalgamation property.

Consider complete lifts $\mathbf{X}, \mathbf{Y}, \mathbf{Z} \in \mathcal{K}_\mathcal{F}$ such that $\mathbf{Z}$ is the substructure induced by both $\mathbf{X}$ and $\mathbf{Y}$ on $Z = X \cap Y$. Put

$$\begin{aligned} \mathbf{A} &= W(\mathbf{X}), \\ \mathbf{B} &= W(\mathbf{Y}), \\ \mathbf{C} &= Sh(\mathbf{Z}). \end{aligned}$$

Consider the free amalgam $\mathbf{D}$ of structures $\mathbf{A}$ and $\mathbf{B}$ over $\mathbf{C}$. Using Lemma 6.5.3, a free amalgam $\mathbf{D}$ is also a witness of $\mathbf{Z}$, $\mathbf{X}$ and $\mathbf{Y}$. To show the amalgamation property holds just take $\mathbf{E} \in \mathcal{K}_\mathcal{F}$ such that $\mathbf{D} \subseteq \mathbf{E}$ with $L(\mathbf{E})$ being complete on $D$. It follows easily that the structure induced by $L(\mathbf{E})$ on $D$ is the amalgamation of $\mathbf{X}$ and $\mathbf{Y}$ over $\mathbf{Z}$.

Consequently there exists an ultrahomogeneous structure $\mathbf{U}'$. Denote the shadow of this structure by $\mathbf{U}$. We show that $\mathbf{U}$ is model complete. This gives existential completeness and $\omega$-saturation of $\mathbf{U}$ as well as the fact that $\mathrm{Aut}(\mathbf{U}) = \mathrm{Aut}(\mathbf{U}')$.

A structure is called model complete if every formula is equivalent to an existential formula. We can make use of fact that $\mathbf{U}'$ is model complete (because every ultrahomogeneous structure is model complete). In this case for any formula $\phi$ in language of $\mathbf{U}$ we can obtain existential formula $\phi'$ in language of $\mathbf{U}'$. What is needed is a transformation of this formula into the language of $\mathbf{U}$. This can be done by replacing every $\vec{v} \in X^i_{\mathbf{U}'}$ by an existential formula testing the presence of a rooted homomorphism $f$ from $\mathcal{P}_i \to \mathbf{U}$ such that $f(\vec{R}_i) = \vec{v}$. Completeness

of witnesses has to be used in order to translate $\vec{v} \notin X^i_{\mathbf{U}'}$ into an existential formula. It follows that the resulting existential formula describes all cases where a homomorphism $\mathcal{P}_i \to \mathbf{U}$ having $f(\vec{R}_i) = \vec{v}$ would produce a homomorphic image of some $\mathbf{F} \in \mathcal{F}$. This means that for every piece $\mathcal{P}_i$ there is a set of rooted structures $\mathcal{P}^1_i, \mathcal{P}^2_i, \ldots, \mathcal{P}^n_i$ such that $\vec{v} \notin X^i_{\mathbf{U}'}$ implies the existence of $m$, $1 \leq m \leq n$ and a rooted homomorphism $g : \mathcal{P}^m_i \to \mathbf{U}$ such that the root is mapped to $\vec{v}$ and moreover the free amalgam of $\mathcal{P}^m_i$ and $\mathcal{P}_i$ is not in $\mathrm{Forb_h}(\mathcal{F})$. □

**Examples.** Theorem 6.5.4 applies to many families $\mathcal{F}$. For example:

1. Let $\mathcal{F}$ be a family of relational trees and $\mathbf{U}$ the canonical universal structure for $\mathrm{Forb_h}(\mathcal{F})$ (if it exists). By Theorem 6.5.4 $\mathrm{rc}(\mathbf{U}) \leq 1$. In fact $\mathbf{U}$ can be seen as a "blown up" core of a homomorphism dual $\mathbf{D}$ (given by [117] even for some infinite families $\mathcal{F}$) where each vertex is replaced by infinitely many vertices and each edge by a random bipartite graph. In this case the bound given by Theorem 6.5.2 is not tight even for $\mathcal{F}$ consisting of an oriented path on 4 vertices.

2. Let $\mathcal{F}_{C_n}$ contain a single odd graph cycle on $n$ vertices. The relational complexity of the canonical universal structure for $\mathrm{Forb_h}(\mathcal{F}_{C_n})$ is at most 2.

3. Let $\mathcal{F}_{odd}$ be the class of all odd graph cycles. The canonical universal structure for $\mathrm{Forb_h}(\mathcal{F}_{odd})$ is the random bipartite graph $\mathbf{B}_2$. By Theorem 6.5.4 we have $\mathrm{rc}(\mathbf{B}) \leq 2$.

### 6.5.2 Lower bounds on relational complexity

We obtain the following bound:

**Theorem 6.5.5.** *Let $\mathcal{F}$ be a finite minimal family of finite connected structures and $\mathbf{U}$ an $\omega$-categorical universal structure for* $\mathrm{Forb_h}(\mathcal{F})$. *Then* $\mathrm{rc}(\mathbf{U})$ *and* $\mathrm{lc}(\mathbf{U})$ *are bounded from below by the size of the largest minimal g-separating g-cut in $\mathcal{F}$.*

We use the following result, proved by a special Ramsey-type construction. This is not a technical finesse, but rather in a way a necessity. It has been shown by [115, 116] that Ramsey classes are related to ultrahomogeneous structures. This connection has been elaborated in the context of topological dynamics in [83].

**Theorem 6.5.6** ([72])**.** *Let $\mathcal{F}$ be a finite minimal family of finite connected relational structures and $\mathcal{K}$ a lift of the class* $\mathrm{Forb_h}(\mathcal{F})$ *adding finitely many new relations of arity at most $r$. If $\mathcal{K}$ contains an ultrahomogeneous lift $\mathbf{U}$ that is universal for $\mathcal{K}$ then the size of minimal g-separating g-cuts of $\mathbf{F} \in \mathcal{F}$ is bounded by $r$.*

Theorem 6.5.5 follows directly as follows. Fix an $\omega$-categorical $\mathbf{U}$ universal for $\mathrm{Forb_h}(\mathcal{F})$. By $\omega$-categoricity of $\mathbf{U}$ we know that $\mathrm{Inv}_k(\mathbf{A})$ is finite and we can now apply Theorem 6.5.6.

**Examples.**

1. Let $\mathcal{F}_P$ consist of the Petersen graph alone and let $\mathbf{U}$ be the canonical structure for $\mathrm{Forb_h}(\mathcal{F}_P)$. Then $\mathrm{rc}(\mathbf{U}) = \mathrm{lc}(\mathbf{U}) = 4$. Recall that the minimal g-separating g-cuts of the Petersen graph are shown in Figure 6.4.

2. The complexity of an $\omega$-categorical graph universal for $\mathrm{Forb_h}(\mathcal{F}_{C_n})$, $n \geq 5$, (graphs without odd cycles of length at most $n$) is at least 2. Combining this with Theorem 6.5.4, we know that the relational complexity of the canonical universal structure for the class $\mathrm{Forb_h}(\mathcal{F}_{C_n})$ is 2. On the other hand, however, this does not hold for the class $\mathcal{F}_{odd}$. We have already shown that the canonical universal graph for the class of all bipartite graphs has relational complexity 2 and lift complexity 1. Finiteness and minimality assumptions are thus both essential in Theorem 6.5.5.

## 6.6 Concluding remarks

The results presented in this chapter can be found in [59, 60], together with several auxiliary results that have been adapted from various sources as referenced. To mention three important sources we can start with the original idea of relational complexity from [29, 27] together with several examples like the homogenization of $L(K_{n,n})$ for $n \geq 4$ and the relational complexity of Johnson and Kneser graphs. The second source [32] treats homogenization, and we have adapted this for the proof of Theorem 6.5.2. The last source is a combination of [73] and particularly [72], where the last-mentioned paper was written with informal input from the present author, and contains a construction used in the proof given of Theorem 6.5.4. In fact a fully detailed proof of this theorem will appear in [58], along with several other results.

The main results of this chapter can be divided into two parts. The first deals with finite graphs. This part is further divided according to the values of relational complexity. For relational complexity equal to 1 one of the results is represented by the relatively simple Proposition 6.3.5, joining the classification of these graphs with the classification of 2-colored $n$-graphs. Proposition 6.3.5 provides a tool for constructing various examples of graphs with relational complexity 1. This is even more interesting because the classification of ultrahomogeneous 2-colored $n$-graphs is still an open problem – see [133]. Moreover the construction underlying this proposition was used to show that any infinite graph having $n$ parts and with lift complexity equal to 1 corresponds to an ultrahomogeneous 2-colored $n$-graph, as stated in Proposition 6.3.7. The main goal here for future work would be the classification of 2-colored $n$-graphs, which would provide insights into graphs with relational complexity 1.

Section 6.3.2, about finite graphs having relational compexity equal 2, starts with the simple but nonetheless important Proposition 6.3.8 showing that all metrically homogeneous graphs have relational complexity 2. This is followed by Proposition 6.3.9 showing by a more technical proof that all trees have relational complexity 2. This section closes with the observation that cographs have the same relational complexity 2.

**Problem 6.6.1.** *Which finite graphs other than metrically homogenous graphs, trees and cographs have relational complexity 2?*

In Section 6.4 we move to a discussion about graphs with high relational complexity. We gave an example of the construction of a graph of arbitrary large finite relational complexity based on a permutation group. Further examples drawn from the work of Cherlin give the relational complexity of well known graph classes such as the Johnson and Kneser graphs. This section closes with Problem 6.4.1, in which we ask for the maximal progression of relational complexity as a function of graph size. Since a complete solution is probably hard it is more reasonable to ask the following roughly defined question.

**Problem 6.6.2.** *Provide examples giving a partial solution to Problem 6.4.1.*

The second part deals with infinite structures. Inspired by Fraïssé's Theorem 1.2.8, for a fixed age $\mathcal{K}$ the problem was to find bounds on the relational complexity of a structure having $\mathcal{K}$ as an age. Since the field of infinite relational structures is quite wide, even when considering the special structure of ages, several additional conditions have to be imposed. At first the restriction to $\omega$-categorical structures has been adopted considering also their characterization via Theorem 1.5.1 by Engeler, Ryll-Nardzewski and Svenonius and Theorem 1.5.2, showing universality of an $\omega$-categorical structure for the class of younger structures. Additionally, only existentially complete structures have been considered so as to overcome the still widely defined range of relational complexity for ages of $\omega$-categorical structures. All necessary properties of structures are encapsulated in the term cannonical universal structure. To conclude this second part we asked the following questions:

I. What is the minimal relational complexity of an $\omega$-categorical structure $\mathbf{U}$ such that $\mathrm{Age}(\mathbf{U}) = \mathcal{K}$?

II. What is the relational complexity of the canonical universal structure of the age $\mathcal{K}$?

The remainder of the chapter provides bounds for relational complexity as well as lift complexity, as summarized in Theorem 6.5.1, although some partial results can be stronger. The first proof of upper bounds for relational complexity is given by Theorem 6.5.2. It makes use of the set of minimal amalgamation failures given by Covington [32]. This seems to be a relatively elegant solution, nevertheless such a set is usually hard to find. For this reason another bound for relational complexity for the canonical universal structure for the age of $\mathrm{Forb}_{\mathrm{h}}(\mathcal{F})$, where $\mathcal{F}$ is a family of connected structures, is given in Theorem 6.5.4, using the size of the largest minimal $g$-saparating g-cut. This proof will appear in [58], although related basic constructions have also been used in [72], a product of informal cooperation among the authors of these two papers. In fact, g-separating g-cuts were also used in proving lower bounds. Theorem 6.5.5 gives lower bounds for relational as well as lift complexity for the $\omega$-categorical structure for $\mathrm{Forb}_{\mathrm{h}}(\mathcal{F})$, where $\mathcal{F}$ is a finite minimal family of connected structures, by using the size of the largest minimal g-separating g-cut.

Both types of complexities are based on ultrahomogeneity. As mentioned in Section 1.7, Cameron and Nešetřil [18] introduced the concept of homomorphism-homogeneous relational structures. It would be interesting to study the topic of this chapter in the context of this type of homogeneity.

# 7. Conclusion

> The game of science is, in principle, without end.
>
> Karl Popper

In this chapter we summarize the conclusions to previous chapters given in Sections 2.6, 3.3, 4.4, 5.3 and 6.6, to which reference should be made for further detail and for corresponding open problems. This chapter retells the whole story that started in the introductory Section 1.10 and continued through subsequent chapters. It serves mainly as an overview and reference list to the particular conclusions we have drawn along the way. Special emphasis is given to the actual contributions of this work and the corresponding papers.

As mentioned in the first chapter, a theme pervading this work is that of symmetric structures, most importantly those defined by having the property that each partial mapping of a predefined type extends to a mapping on the whole structure, again of a given type. Examples of symmetric structures defined in this way include those that are ultrahomogeneous (partial isomorphisms extend to isomorphisms) and homomorphism-homogeneous (partial homomorphisms extend to homomorphisms). These properties of structures are a focus of this work. The theoretical background required has drawn from various mathematical disciplines such as graph theory, combinatorics, group theory, model theory and others. These theoretical essentials are described in Chapter 1, where most of the theory needed for handling these structures is introduced.

Also in the introductory chapter, the idea of structure complexity, which is the main topic of this work, is given a preliminary introduction. Roughly speaking, the structure complexity is lower the "nearer" the structure is to meeting the conditions of the version of homogeneity that is under study. The underlying language is a determinative variable here, therefore the main task proves to be studying homogeneity of the corresponding class of structures when changing the definition of this underlying language. The distance which determines proximity to homogeneity is represented by the arities of relations used to extend the original language.

There are basically two ways to proceed. The first is to study how the classification of homogeneous structures changes when we change the underlying language of relational structures by adding relations of relatively small arity. The corresponding results, summarized in Sections 7.2 and 7.3 below, provide a classification of homomorphism-homogeneous $L$-colored graphs for various $L$.

The second way, described in Section 7.4 below, uses another approach represented by relational and lift complexity. In this case for a given structure we search for a minimal arity $k$ such that there exists a set of relations each of arity at most $k$ such that the original structure extended using these relations is ultrahomogeneous. This task is also considered more generally for a class of structures via the analysis of the corresponding universal structure.

Infinite countable universal structures meeting the homogeneity condition have a surprising connection with graph limits and consequently with complex networks [100, 118]. This connection together with the corresponding studies of

complex networks are described in Section 7.1.

## 7.1 Homogeneous structures as motivation for complex networks

One of the connections between complex networks and homogeneous structures finds realization in graphs limits as defined by Lovász [100] and Nešetřil and Ossona de Mendez [118]. This connection is described in Section 2.6. The notion of $FO$-convergence [118] can represent limits of graphs as defined either by Lovász and Szegedy [101] or by Benjamini and Schramm [8]. For example, for $\omega$-categorical structures it is seen that for a sequence of graphs having an ultrahomogeneous limit the corresponding $FO$-convergence in fact reduces to $QF$-convergence. Considering this together with the connection of graph limits to complex networks shown by Lovász [100] suggests that highly symmetric structures may occur naturally as limits of graph sequences representing complex networks. The question was raised in Problem 2.6.3 of whether there is a class of complex networks which can be well represented by a network growth model such that this model is $FO$-convergent and has an $\omega$-categorial structure as its limit.

To become more familiar with complex networks requires becoming familiar with how they are constructed. For this reason several studies have been made by the author in order to clarify this process. A list of these studies follows.

The first area concerns the brain, analysed by the method of functional magnetic resonance imaging (fMRI). The problem studied was to understand the effect of nonlinearity in time series representing underlying subsystems in the construction of complex networks and therefore whether the measure that is used to define edges should be the correlation coefficient or mutual information. The results show that the differences resulting from which choice is made are fairly negligible. The methods and results are presented in Section 2.3 and can be found in the corresponding paper [56].

Inspired by the case of brain networks, a similar study has been carried out for climate networks, which are quite different from various points of view. However, even for this type of system it has been shown that observed nonlinearity effects can be corrected by proper preprocessing. These results are described in Section 2.4.2 and details can be found in [64]. In the same section other results are given that correct the bias in computation of association measures caused by the dynamics of the corresponding time series by introducing Z-scores based on independent Fourier transform surrogate data [128].

Additionally for climate networks, we have carried out an analysis of the reliability of causal measures, for example conditional mutual information, for more complicated directed networks. This analysis is described in Section 2.4.3 and details can be found in [65]. Finally, small-world phenomena, introduced at the beginning of Chapter 2, are analysed in Section 2.5. It has been shown that even for models generated from completely random networks the small-world characteristic indicates the presence of the phenomenon and thus its reliability is questionable, especially considering the transitive character of the corresponding correlation measure [63].

## 7.2 Colored graphs

The first step in studying the complexity of structures is through extending their language and making a classification. Since the classification of countable ultrahomogeneous graphs has been completed, a natural candidate is the corresponding problem for homomorphism-homogeneity. For finite graphs classification is easy and therefore the starting point is finite structures with an extended language – namely bicolored graphs. This also follows the suggestion of Macpherson in [104] that one possible future step would be to classify homomorphism-homogeneous graphs with edges colored using two or more colors. The classification of homomorphism-homogeneous structures of this type is given by Theorem 3.2.5. This result appears in [61].

The importance of this proof lies mainly in introducing the notion called the pumping argument that is used later in the classification of $L$-colored homomophism-homogeneous graphs. In the concluding Section 3.3 an open problem concerning coincidence of the classes **HH** and **MH** is discussed. It turns out that arguments from this chapter cannot be used to resolve this problem completely for the class of bicolored graphs. However, as mentioned there, using results from the following chapter about homomorphism-homogeneous $L$-colored graphs it can be shown that these classes in fact coincide for bicolored graphs. In Section 3.3 other open questions and problems are posed, such as the problem of classifing countably infinite homomophism-homogeneous bigraphs (Problem 3.3.2), whether the classes **HH** and **MH** coincide for countably infinite bicolored graphs (Problem 3.3.3), and whether or not the classifications of homomorphism-homogeneous bicolored graphs and 3-edge-colored graphs differ (Problem 3.3.4).

The classification of bicolored graphs is followed by the classification of $L$-colored graphs in Chapter 4. This topic is represented by two consecutive papers [61, 57] together covering the content of the chapter as a whole. The classification is given by two theorems. Firstly, Theorem 4.2.3 provides the classification of finite homomorphism-homogeneous and monomorphism-homogeneous $L$-colored graphs when $L$ is a chain. Secondly, the classification with same parameters except that $L$ is assumed to be a diamond is given by Theorem 4.3.4. In order to prove these results a generalization of the pumping argument has been used.

As a side effect of the work in Chapter 4, examples of $L$-colored graphs that show the classes **HH** and **MH** do not coincide for these structures emerged for both the finite and infinite case. In Section 4.4 there are detailed conclusions for this chapter, in which several open questions and conjectures arise, such as the problem of classifying all finite homomorphism-homogeneous $L$-colored graphs for general $L$ (Problem 4.4.1) and whether the classes **MH** and **HH** coincide for vertex-uniform $L$-colored graphs (Conjecture 4.4.1). Moreover if Conjecture 4.4.1 were true, one could then ask about the coincidence of these classes for countably infinite vertex-uniform graphs, as conditionally posed in Problem 4.4.2.

## 7.3 Morphism extension classes

The question of the coincidence of classes **HH** and **MH** motivates a wider exploration of the whole category of classes defined by various types of homogeneity

and mutual coincidence. As the main result of Chapter 5, the class of $L$-colored graphs is shown to be sufficient to have all mutual relationships hold without actual coincidence.

In fact even finite $L$-colored graphs would be enough, were it not for the fact that the mutual coincidence of the classes **IM** and **II** can only be removed by introducing countably infinite $L$-colored graphs.

This chapter largely consists of a collection of results from other authors and other chapters of this thesis put into a broader context. It closes with asking about the morphism extension classes hierarchy for $L$-colored graphs when using a connected-homomorphism-homogeneous variant (Problem 5.3.1).

## 7.4 Relational complexity

Complementary to the analysis in Chapter 4 of homomorphism-homogeneous $L$-colored graphs, Chapter 6 considers the language of a relational structure as a variable and searches for its "minimal extension" to obtain an ultrahomogeneous lift. This process is represented by the notions of relational and lift complexity. The results of this chapter can be found in two consecutive papers [60, 59] which together provide an analysis of these complexities. For more detailed conclusions see Section 6.6.

The analysis is divided into two parts. The first part deals with finite graphs, further subdivided according to complexity values. The main result for complexity value 1 is given by Proposition 6.3.5, joining the classification of graphs with relational complexity 1 with the classification of 2-colored $k$-graphs, which is even more interesting since their classification is still an open problem. On the other hand, for relational complexity 2 there are several propositions establishing familes of graphs with relational complexity 2, namely Proposition 6.3.8 for metrically homogeneous graphs, Proposition 6.3.9 for trees and Proposition 6.3.10 for cographs. This analysis closes with several open problems, for example Problem 6.6.1 asking for a complete classification of graphs with relational complexity 2 or Problem 6.4.1 asking for determining the largest relational complexity a graph with size $n$ can have.

The second part deals with countably infinite structures. Namely it searches for bounds on the relational complexity of universal structures for given classes of relational structures. The class that is of particular interest in this work is $\mathrm{Forb}_{\mathrm{h}}(\mathcal{F})$ for various $\mathcal{F}$. The overall, although not the most general, result of this part is given in Theorem 6.5.1, providing a lower bound on the relational complexity of an $\omega$-categorical universal structure $U$ for the class $\mathrm{Forb}_{\mathrm{h}}(\mathcal{F})$. When $\mathcal{F}$ is a finite minimal family of finite connected relational structures the relational complexity of $U$ is bounded below by the size of the largest minimal g-separating g-cut. Moreover if $U$ is a canonical universal structure for such a class the size of the largest minimal g-separating g-cut gives the exact value of the relational complexity.

The mentioned lower bound is given by Theorem 6.5.5 using a simple observation in the light of other results from Hubička and Nešetřil [72], namely Theorem 6.5.6. The upper bound is provided mainly in Theorem 6.5.4, where the construction of a specific lift-preserving automorphism group is given and it is proved that the relational complexity of the canonical universal structure for the

class $\mathrm{Forb}_{\mathrm{h}}(\mathcal{F})$ is bounded above by the size of the largest minimal g-separating g-cut. The assumptions about $\mathcal{F}$ make the class $\mathrm{Forb}_{\mathrm{h}}(\mathcal{F})$ in fact wider than in Theorem 6.5.1, however there is a condition for a universal structure to be in fact a canonical universal structure. This ensures, mainly through existential completeness, that the resulting lift does not modify the automorphism group. Additionally, there is a simple proof using the notion of amalgamation failure to bound relational complexity, as stated in Theorem 6.5.2, which is rather an alternative to the previously mentioned bound, since g-separating g-cuts are more easily handled.

In the concluding Section 6.6 there is a summary of results and the formulation of several open problems.

# Relevant papers of the author

Below we highlight relevant publications of the author, all of which are also given in the bibliography, where they are used for actual referencing from the text.

## Homogeneous relational structures

**D. Hartman**, J. Hubička, and D. Mašulović. Homomorphism-homogeneous $L$-colored graphs. *European Journal of Combinatorics*, 35(0):313–323, 2014.

**D. Hartman**, J. Hubička, and J. Nešetřil. Combinatorial bounds on relational complexity. In J. Nešetřil and M. Pellegrini, editors, *The Seventh European Conference on Combinatorics, Graph Theory and Applications*, volume 16 of *CRM Series*, pages 573–578. Scuola Normale Superiore, 2013.

**D. Hartman** and D. Mašulović. Towards finite homomorphism-homogeneous relational structures. *Electronic Notes in Discrete Mathematics*, 38(0):443 – 448, 2011.

**D. Hartman**, J. Hubička, and J. Nešetřil. Complexities of relational structure. *accepted for publication in Mathematica Slovaca*, arXiv:1309.4266, 2014.

## Complex networks

**D. Hartman**, J. Hlinka, M. Paluš, D. Mantini, and M. Corbetta. The role of nonlinearity in computing graph-theoretical properties of resting-state functional magnetic resonance imaging brain networks. *Chaos*, 21(1):art.num.013119, 2011.

J. Hlinka, **D. Hartman**, and M. Paluš. Small-world topology of functional connectivity in randomly connected dynamical systems. *Chaos*, 22(3), 2012.

J. Hlinka, **D. Hartman**, M. Vejmelka, D. Novotná, and M. Paluš. Non-linear dependence and teleconnections in climate data: sources, relevance, nonstationarity. *Climate Dynamics*, pages 1–14, 2013.

J. Hlinka, **D. Hartman**, M. Vejmelka, J. Runge, N. Marwan, J. Kurths, and M. Paluš. Reliability of Inference of Directed Climate Networks Using Conditional Mutual Information. *Entropy*, 15(6):2023–2045, 2013.

# References


[1] R. Albert, H. Jeong, and A.L. Barabasi. Internet - Diameter of the World-Wide Web. *Nature*, 401(6749):130–131, 1999.

[2] A. Arenas, A. Diaz-Guilera, J. Kurths, Y. Moreno, and C. Zhou. Synchronization in complex networks. *Physics Report - Review Section of Physics Letters*, 469(3):93–153, 2008.

[3] A. L. Barabasi and R. Albert. Emergence of scaling in random networks. *Science*, 286(5439):509–512, 1999.

[4] A. L. Barabasi and Z. N. Oltvai. Network biology: Understanding the cell's functional organization. *Nature Review Genetics*, 5(2):101–U15, 2004.

[5] A. G. Barnston and R. E. Livezey. Classification, seasonality and persistence of low-frequency atmorpheric circulation patterns. *Monthly Weather Review*, 115(6):1083–1126, 1987.

[6] D. S. Bassett and E. Bullmore. Small-world brain networks. *The Neuroscientist*, 12(6):512–523, 2006.

[7] D. S. Bassett and E. Bullmore. Human brain networks in health and disease. *Current Opinion in Neurology*, 22(4):340–347, 2009.

[8] I. Benjamini and O. Schramm. Recurrence of distributional limits of finite planar graphs. *Electronic Journal of Probabability*, 6:1–13, 2001, n. 23.

[9] N. Berger, C. Borgs, J. T. Chayes, and A. Saberi. Asymptotic behavior and distributional limits of preferential attachment graphs. *Annals of Probability*, 42(1):1–40, 2014.

[10] T. Berners-Lee, W. Hall, J. Hendler, N. Shadbolt, and D. J. Weitzner. Creating a science of the Web. *Science*, 313(5788):769–771, 2006.

[11] S. Bialonski, M.-T. Horstmann, and K. Lehnertz. From brain to earth and climate systems: Small-world interaction networks or not? *Chaos*, 20(1):art.num.013134, 2010.

[12] S. Bialonski, M. Wendler, and K. Lehnertz. Unraveling Spurious Properties of Interaction Networks with Tailored Random Networks. *PLOS ONE*, 6(8):art.num.e22826, 2011.

[13] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D.-U. Hwang. Complex networks: Structure and dynamics. *Physics Reports*, 424(4-5):175 – 308, 2006.

[14] R. L. Buckner, J. Sepulcre, T. Talukdar, F. M. Krienen, H. Liu, T. Hedden, J. R. Andrews-Hanna, R. A. Sperling, and K. A. Johnson. Cortical Hubs Revealed by Intrinsic Functional Connectivity: Mapping, Assessment of Stability, and Relation to Alzheimer's Disease. *Journal of Neuroscience*, 29(6):1860–1873, 2009.

[15] E. Bullmore and O. Sporns. Complex brain networks: graph theoretical analysis of structural and functional systems. *Nature Reviews Neuroscience*, 10(3):186–198, 2009.

[16] P. J. Cameron. The Random Graph. In Graham, R. L. and Nešetřil, J. and Butler, S., editor, *The Mathematics of Paul Erdős II*, pages 353–378. Springer, 2nd edition, 2013.

[17] P. J. Cameron and D. C. Lockett. Posets, homomorphisms and homogeneity. *Discrete Mathematics*, 310(3):604–613, 2010.

[18] P. J. Cameron and J. Nešetřil. Homomorphism-homogeneous relational structures. *Combinatorics probability & computing*, 15(1-2):91–103, 2006.

[19] P.J. Cameron. The age of a relational structure. *Discrete Mathematics*, 95(1-3):49–67, 1991.

[20] G. Cantor. Beiträge zur begründing der transfiniten mengerlehre. *Mathematische Annalen*, 46(4):481–512, 1895.

[21] A. Cardillo, A. Scellato, V. Latora, and S. Porta. Structural properties of planar graphs of urban street patterns. *Physical Review E*, 73(6, Part 2):art.num.066107, 2006.

[22] G. Cherlin. Sporadic homogeneous structures. In *The Gelfand Mathematical Seminars (1994-1999)*. Birkhäuser Boston, 2000.

[23] G. Cherlin. Two problems on homogeneous structures, revisited. In Grohe, M. and Makowsky, J. A., editor, *Model Theoretic Methods in Finite Combinatorics*, volume 558 of *Contemporary Mathematics*, pages 319–415, 2011.

[24] G. Cherlin and S. Shelah. Universal graphs with a forbidden subtree. *Journal of Combinatorial Theory, Series B*, 97(3):293 – 333, 2007.

[25] G. Cherlin, S. Shelah, and N. D. Shi. Universal graphs with forbidden subgraphs and algebraic closure. *Advances in Applied Mathematics*, 22(4):454–491, 1999.

[26] G. I. Cherlin. The classification of countable homogeneous directed graphs and countable homogeneous $n$-tournaments. *Memoirs of American Mathematical Society*, 131(621):161, 1998.

[27] G. I. Cherlin. On the relational complexity of a finite permutation group. *submitted*, `http://www.math.rutgers.edu/~cherlin/Paper/index.html`, 2013.

[28] G. L. Cherlin and P. Komjáth. There is no universal countable pentagon-free graph. *Journal of Graph Theory*, 18(4):337–341, 1994.

[29] G. L. Cherlin, G. A. Martin, and D. H. Saracino. Arities of permutation groups: Wreath products and k-sets. *Journal of Combinatorial Theory Series A*, 74(2):249–286, 1996.

[30] R. Cohen, K. Erez, D. ben Avraham, and S. Havlin. Resilience of the Internet to random breakdowns. *Physical Review Letters*, 85(21):4626–4628, 2000.

[31] T. M. Cover and J. A. Thomas. *Elements of Information Theory*. Wiley, 2012.

[32] J. Covington. Homogenizable Relational Structures. *Illinois J. Mathematics*, 34(4):731–743, 1990.

[33] M. Ding, Y. Chen, and S. L. Bressler. Granger causality: Basic theory and application to neuroscience. In *Handbook of Time Series Analysis*, pages 437–460. Wiley-VCH Verlag GmbH and Co. KGaA, 2006.

[34] J. F. Donges, Y. Zou, N. Marwan, and J. Kurths. The backbone of the climate network. *EPL (Europhysics Letters)*, 87(4):art.num.48007, 2009.

[35] K. T. D. Eames and M. J. Keeling. Modeling dynamic and network heterogeneities in the spread of sexually transmitted diseases. *Proceedings of the National Academy of Sciences of the United States of America*, 99(20):13330–13335, 2002.

[36] E. Engeler. Äquivalenzklassen von $n$-tupeln. *Mathematical Logic Quarterly*, 5(14-24):340–345, 1959.

[37] H. Enomoto. Combinatorially homogeneous graphs. *Journal of Combinatorial Theory, Series B*, 30(2):215–223, 1981.

[38] P. Erdős and A. Rényi. On random graphs. *Publicationes Mathematicae Debrecen*, 6:290–297, 1959.

[39] P. Erdős and A. Rényi. On the evolution of random graphs. *Publications of the Mathematical Institute of the Hungarian Academy of Sciences*, 5:17–61, 1960.

[40] P. Erdős and A. Rényi. Asymmetric graphs. *Acta Mathematica Academiae Scientiarum Hungaricae*, 14:295 – 315, 1963.

[41] M. G. Everett. Role similarity and complexity in social networks. *Social Networks*, 7(4):353–359, 1985.

[42] M. E. Falagas, E. I. Pitsouni, G. A. Malietzis, and G. Pappas. Comparison of PubMed, Scopus, Web of Science, and Google Scholar: strengths and weaknesses. *Faseb Journal*, 22(2):338–342, 2008.

[43] M. Faloutsos, P. Faloutsos, and C. Faloutsos. On power-law relationships of the Internet topology. In Guerin, R., editor, *ACM SIGCOMM'99 Conference: Applications, Technologies, Architectures, and Protocols for Computer Communications*, volume 29 of *Computer Communications Review*, pages 251–262, 1999.

[44] R. Fraïssé. Sur certains relations qui généralisent l'ordre des nombres rationnels. *Comptes Rendus de l'Académie des Sciences*, 15:91–103, 2006.

[45] L. C. Freeman. Set of measures of centrality based on betweenness. *Sociometry*, 40(1):35–41, 1977.

[46] L. C. Freeman. Centrality in social networks conceptual clarification. *Social Networks*, 1(3):215–239, 1979.

[47] K. J. Friston. Functional and effective connectivity in neuroimaging: A synthesis. *Human Brain Mapping*, 2(1-2):56–78, 1994.

[48] K. J. Friston. Beyond phrenology: What can neuroimaging tell us about distributed circuitry? *Annual Review of Neuroscience*, 25:221–250, 2002.

[49] K. J. Friston. *Statistical Parametric Mapping: The Analysis of Functional Brain Images*. Academic Press. Elsevier/Academic Press, 2007.

[50] A. Gardiner. Homogeneous graphs. *Journal of Combinatorial Theory Series B*, 20(1):94–102, 1976.

[51] A. Gardiner. Homogeneity conditions in graphs. *Journal of Combinatorial Theory Series B*, 24(3):301–310, 1978.

[52] J. F. Geweke. Measurement of linear dependence and feedback between multiple time series. *Journal of the Americal Statistical Association*, 77(378):304–313, 1982.

[53] M. Girvan and M. E. J. Newman. Community structure in social and biological networks. *Proceedings of the National Academy of Sciences of the United States of America*, 99(12):7821–7826, 2002.

[54] C. W. J. Granger. Investigating causal relations by econometric models and cross-spectral methods. *Econometrica*, 37(3):424–438, 1969.

[55] R. Guimera, S. Mossa, A. Turtschi, and L. A. N. Amaral. The worldwide air transportation network: Anomalous centrality, community structure, and cities' global roles. *Proceedings of the National Academy of Sciences of the United States of America*, 102(22):7794–7799, 2005.

[56] D. Hartman, J. Hlinka, M. Paluš, D. Mantini, and M. Corbetta. The role of nonlinearity in computing graph-theoretical properties of resting-state functional magnetic resonance imaging brain networks. *Chaos*, 21(1):art.num.013119, 2011.

[57] D. Hartman, J. Hubička, and D. Mašulović. Homomorphism-homogeneous $L$-colored graphs. *European Journal of Combinatorics*, 35(0):313–323, 2014.

[58] D. Hartman, J. Hubička, and J. Nešetřil. On classes of relational structures with bounded relational complexity. *in prep.*

[59] D. Hartman, J. Hubička, and J. Nešetřil. Combinatorial bounds on relational complexity. In J. Nešetřil and M. Pellegrini, editors, *The Seventh European Conference on Combinatorics, Graph Theory and Applications*, volume 16 of *CRM Series*, pages 573–578. Scuola Normale Superiore, 2013.

[60] D. Hartman, J. Hubička, and J. Nešetřil. Complexities of relational structure. *accepted for publication in Mathematica Slovaca*, preprint at `http://arxiv.org/abs/1309.4266`, 2014.

[61] D. Hartman and D. Mašulović. Towards finite homomorphism-homogeneous relational structures. *Electronic Notes in Discrete Mathematics*, 38(0):443 – 448, 2011.

[62] P. Hell and J. Nešetřil. *Graphs and Homomorphisms*. Oxford Lecture Series in Mathematics and Its Applications. OUP Oxford, 2004.

[63] J. Hlinka, D. Hartman, and M. Paluš. Small-world topology of functional connectivity in randomly connected dynamical systems. *Chaos*, 22(3):art.num.033107, 2012.

[64] J. Hlinka, D. Hartman, M. Vejmelka, D. Novotná, and M. Paluš. Nonlinear dependence and teleconnections in climate data: sources, relevance, nonstationarity. *Climate Dynamics*, 42(7-8):1873–1886, 2013.

[65] J. Hlinka, D. Hartman, M. Vejmelka, J. Runge, N. Marwan, J. Kurths, and M. Paluš. Reliability of Inference of Directed Climate Networks Using Conditional Mutual Information. *Entropy*, 15(6):2023–2045, 2013.

[66] J. Hlinka, M. Paluš, M. Vejmelka, D. Mantini, and M. Corbetta. Functional connectivity in resting-state fMRI: Is linear correlation sufficient? *Neuroimage*, 54(3):2218–2225, 2011.

[67] W. Hodges. *A Shorter Model Theory*. Cambridge University Press, 1997.

[68] P. Holme. Detecting degree symmetries in networks. *Physical Review E*, 74(3, 2):art.num.036107, 2006.

[69] P. Holme, B. J. Kim, C. N. Yoon, and S. K. Han. Attack vulnerability of complex networks. *Physical Review E*, 65(5, 2):art.num.056109, 2002.

[70] P. Holme and J. Saramäki. Temporal networks. *Physics Report - Review Section of Physics Letters*, 519(3):97–125, 2012.

[71] B Horwitz. The elusive concept of brain connectivity. *Neuroimage*, 19(2, 1):466–470, 2003.

[72] J. Hubička and J. Nešetřil. Homomorphism and embedding universal structures for restricted classes. in preparation, preprint at `http://arxiv.org/abs/0909.4939`.

[73] J. Hubička and J. Nešetřil. Universal structures with forbidden homomorphisms. In *J. Väänänen Festschrift*. De Gruyter, to appear, preprint at `http://arxiv.org/abs/0907.4079`.

[74] M. D. Humphries and K. Gurney. Network 'Small-World-Ness': A Quantitative Method for Determining Canonical Network Equivalence. *PLOS ONE*, 3(4):art. num. e0002051, 2008.

[75] E. V. Huntington. The continuum as a type of order: an exposition of the model theory. *Annals of Mathematics*, 6:178 – 179, 1904.

[76] A. Ilić, D. Mašulović, and Uroš Rajković. Finite homomorphism-homogeneous tournaments with loops. *Journal of Graph Theory*, 59(1):45–58, 2008.

[77] B. J. Jansen, M. Zhang, K. Sobel, and A. Chowdury. Twitter Power: Tweets as Electronic Word of Mouth. *Journal of the American Society for Information Science and Technology*, 60(11):2169–2188, 2009.

[78] T. Jenkinson, J. K. Truss, and D. Seidel. Countable homogeneous multipartite graphs. *European Journal of Combinatorics*, 33(1):82–109, 2012.

[79] H. Jeong, S. P. Mason, A. L. Barabasi, and Z. N. Oltvai. Lethality and centrality in protein networks. *Nature*, 411(6833):41–42, 2001.

[80] P. W. Jones. First- and second-order conservative remapping schemes for grids in spherical coordinates. *Monthly Weather Review*, 127(9):2204–2210, 1999.

[81] E. Kalnay, M. Kanamitsu, R. Kistler, W. Collins, D. Deaven, L. Gandin, M. Iredell, S. Saha, G. White, J. Woollen, Y. Zhu, M. Chelliah, W. Ebisuzaki, W. Higgins, J. Janowiak, K. C. Mo, C. Ropelewski, J. Wang, A. Leetmaa, R. Reynolds, R. Jenne, and D. Joseph. The NCEP/NCAR 40-year reanalysis project. *Bulletin of the American Meteorological Society*, 77(3):437–471, 1996.

[82] F. Karinthy. Láncszemek. In *Minden másképpen van*. Budapest: Athenaeum, 1929.

[83] A. S. Kechris, V. G. Pestov, and S. Todorčević. Fraïssé limits, Ramsey theory, and topological dynamics of automorphism groups. *Geom. Funct. Anal.*, 15(1):106–189, 2005.

[84] M. J. Keeling and K. T. D. Eames. Networks and epidemic models. *Journal of the Royal Society Interface*, 2(4):295–307, 2005.

[85] R. Klein. *Concrete and Abstract Voronoi Diagrams*. Lecture Notes in Computer Science. Springer, 1989.

[86] J. Knight and A. H. Lachlan. Shrinking, stretching, and codes for homogeneous structures. *Lecture Notes in Mathematics*, 1292:192–229, 1987.

[87] P. Komjath, A. H. Mekler, and J. Pach. Some Universal Graphs. *Israel Journal of Mathematics*, 64(2):158–168, 1988.

[88] T. Kreuz, F. Mormann, R. G. Andrzejak, A. Kraskov, K. Lehnertz, and P. Grassberger. Measuring synchronization in coupled model systems: A comparison of different approaches. *Physica D - Nonlinear Phenomena*, 225(1):29–42, 2007.

[89] K. Kutnar and D. Marusic. Recent Trends and Future Directions in Vertex-Transitive Graphs. *Ars Mathematica Contemporanea*, 1(2):112–125, 2008.

[90] J. E. Kutzbach. Large-scale features of monthly mean northen hemisphere anomaly maps of sea-level pressure. *Monthly Weather Review*, 98(9):708–716, 1970.

[91] A. H. Lachlan. Countable homogeneous tournaments. *Transactions of the American Mathematical Society*, 284(2):431–461, 1984.

[92] A. H. Lachlan and R. E. Woodrow. Countable ultrahomogeneous graphs. *Transactions of the American Mathematical Society*, 262(1):51–94, 1980.

[93] E. Langford, N. Schwertman, and M. Owens. Is the property of being positively correlated transitive? *American Statistician*, 55(4):322–325, 2001.

[94] V. Latora and M. Marchiori. Efficient behavior of small-world networks. *Physical Review Letters*, 87(19):art.num.198701, 2001.

[95] K. Lewis, J. Kaufman, M. Gonzalez, A. Wimmer, and N. Christakis. Tastes, ties, and time: A new social network dataset using facebook.com. *Social Networks*, 30(4):330 – 342, 2008.

[96] W. Liao, Z. Zhang, Z. Pan, D. Mantini, J. Ding, X. Duan, C. Luo, G. Lu, and H. Chen. Altered Functional Connectivity and Small-World in Mesial Temporal Lobe Epilepsy. *PLOS ONE*, 5(1):art.num.e8525, 2010.

[97] D. C. Lockett. Connected-homomorphism-homogeneous graphs. *pre-print*, `http://arxiv.org/abs/1207.4373`, 2012.

[98] D. C. Lockett and J. K. Truss. Some more notions of homomorphism-homogeneity. *manuscript still in preparation*, preprint accessible at `http://www1.maths.leeds.ac.uk/pure/staff/truss/debbie3.pdf`, 2012.

[99] F. Lorrain and H. C. White. Structural equivalence of individuals in social networks. *The Journal of Mathematical Sociology*, 1(1):49–80, 1971.

[100] L. Lovász. *Large Networks and Graph Limits*. American Mathematical Society colloquium publications. American Mathematical Society, 2012.

[101] L. Lovász and B. Szegedy. Limits of dense graph sequences. *Journal of Combinatorial Theory, Series B*, 96(6):933 – 957, 2006.

[102] H. Lütkepohl. *New Introduction to Multiple Time Series Analysis*. Springer, 2006.

[103] B. D. MacArthur, R. J. Sanchez-Garcia, and J. W. Anderson. Symmetry in complex networks. *Discrete Applied Mathematics*, 156(18):3525–3531, 2008.

[104] D. Macpherson. A survey of homogeneous structures. *Discrete Mathematics*, 311(15):1599 – 1634, 2011.

[105] N. Malik, B. Bookhagen, N. Marwan, and J. Kurths. Analysis of spatial and temporal extreme monsoonal rainfall over South Asia using complex networks. *Climate Dynamics*, 39(3-4):971–987, 2012.

[106] R. N. Mantegna. Hierarchical structure in financial markets. *European Physical Journal B*, 11(1):193–197, 1999.

[107] S. Maslov and K. Sneppen. Specificity and stability in topology of protein networks. *Science*, 296(5569):910–913, 2002.

[108] J Matoušek and J. Nešetřil. *Invitation to discrete mathematics*. Oxford University Press, Oxford, 1998.

[109] D. Mašulović. Homomorphism-homogeneous partially ordered sets. *Order*, 24(4):215–226, 2007.

[110] A. R. McIntosh. Towards a network theory of cognition. *Neural Networks*, 13(8-9):861–870, 2000.

[111] M. J. McPhaden, S. E. Zebiak, and M. H. Glantz. ENSO as an integrating concept in Earth science. *Science*, 314(5806):1740–1745, 2006.

[112] G. S. Medvedev. Small-world networks of kuramoto oscillators. *Physica D: Nonlinear Phenomena*, 266(0):13 – 22, 2014.

[113] L. A. Meyers, B. Pourbohloul, M. E. J. Newman, D. M. Skowronski, and R. C. Brunham. Network theory and SARS: predicting outbreak diversity. *Journal of Theoretical Biology*, 232(1):71–81, 2005.

[114] S. Milgram. Small-world problem. *Phychology Today*, 1(1):61–67, 1967.

[115] J. Nešetřil. For graphs there are only 4 types of hereditary Ramsey Classes. *J. Combin. Theory B*, 46(2):127–132, 1989.

[116] J. Nešetřil. Ramsey classes and homogeneous structures. *Combinatorics, Probablity and Computing*, 14(1-2):171–189, 2005.

[117] J. Nešetřil and C. Tardif. Duality theorems for finite structures (characterising gaps and good characterisations). *Journal of Combinatorial Theory, Series B*, 80(1):80–97, 2000.

[118] J. Nešetřil and P. Ossona de Mendez. A unified approach to structural limits, and limits of graphs with bounded tree-depth. *preprint*, `http://arxiv.org/abs/1303.6471`.

[119] M. E. J. Newman. The structure of scientific collaboration networks. *Proceedings of the National Academy of Sciences of the United States of America*, 98(2):404–409, 2001.

[120] M. E. J. Newman. Assortative mixing in networks. *Physical Review Letters*, 89(20):art.num.208701, 2002.

[121] M. E. J. Newman. The structure and function of complex networks. *SIAM Review*, 45(2):167–256, 2003.

[122] M. E. J. Newman. A measure of betweenness centrality based on random walks. *Social Networks*, 27(1):39–54, 2005.

[123] M. E. J. Newman, S. H. Strogatz, and D. J. Watts. Random graphs with arbitrary degree distributions and their applications. *Physical Review E*, 64(2, 2):art.num.026118, 2001.

[124] S. Ogawa, T. M. Lee, A. R. Kay, and D. W. Tank. Brain magnetic-resonance-imaging with contrast dependent on blood oxygenation. *Proceedings of the National Academy of Sciences of the United States of America*, 87(24):9868–9872, 1990.

[125] J. P. Onnela, A. Chakraborti, K. Kaski, J. Kertesz, and A. Kanto. Dynamics of market correlations: Taxonomy and portfolio analysis. *Physical Review E*, 68(5, Part 2):art.num.056110, 2003.

[126] M. Paluš. Detecting phase synchronization in noisy systems. *Physics Letters A*, 235(4):341–351, 1997.

[127] M. Paluš, V. Albrecht, and I. Dvořák. Information theoretic test for nonlinearity in time-series. *Physics Letters A*, 175(3-4):203–209, 1993.

[128] M. Paluš, D. Hartman, J. Hlinka, and M. Vejmelka. Discerning connectivity from dynamics in climate networks. *Nonlinear Process Geophysics*, 18(5):751–763, 2011.

[129] C. Pech and M. Pech. Constraint Satisfaction with Weakly Oligomorphic Template. *manuscript submitted*, preprint accessible at `http://sites.dmi.rs/personal/pechm/AgeHHnew.pdf`, 2011.

[130] D. Prichard and J. Theiler. Generating surrogate data for time-series with several simultaneously measured variables. *Physical Review Letters*, 73(7):951–954, 1994.

[131] R. Rado. Universal graphs and universal functions. *Acta Arithmetica*, 9(3):2, 1964.

[132] S. Redner. How popular is your paper? an empirical study of the citation distribution. *The European Physical Journal B*, 4(2):131–134, 1998.

[133] S. E. Rose. *Classification of Countable Homogeneous* 2*-Graphs*. PhD thesis, University of Leeds, Leeds, England, 2011. reprint available at `http://etheses.whiterose.ac.uk/1750/`.

[134] M. Rosvall, A. Trusina, P. Minnhagen, and K. Sneppen. Networks and cities: An information perspective. *Physical Review Letters*, 94(2):art.num.028701, 2005.

[135] M. Rubinov and O. Sporns. Weight-conserving characterization of complex functional brain networks. *NeuroImage*, 56(4):2068 – 2079, 2011.

[136] M. Rusinov and P. Schweitzer. Homomorphism–homogeneous graphs. *Journal of Graph Theory*, 65(3), 2010.

[137] C. Ryll-Nardzewski. On the categoricity in power $\leq \aleph_0$. *Bull. Acad. Polon. Sci. Math. Astr. Phys.*, 7:545–548, 1959.

[138] A. H. Schmerl. Countable homogeneous partially ordered sets. *Algebra Universalis*, 9(3):317–321, 1979.

[139] S. Schnettler. A small world on feet of clay? a comparison of empirical small-world studies against best-practice criteria. *Social Networks*, 31(3):179 – 189, 2009.

[140] S. Schnettler. A structured overview of 50 years of small-world research. *Social Networks*, 31(3):165 – 178, 2009.

[141] M. A. Serrano and M. Boguna. Topology of the world trade web. *Physical Review E*, 68(1, Part 2):art.num.015101, 2003.

[142] J. Sheehan. Smoothly embeddable subgraphs. *Journal of the London mathematical society – second series*, 9:212–218, 1974.

[143] N. Svenonius. $\aleph_0$-categoricity in first-order predicate calculus. *Theoria*, 25:82–94, 1955.

[144] L. W. Swanson. *Brain Architecture*. Oxford University Press, 2003.

[145] A. A. Tsonis and P. J. Roebber. The architecture of the climate network. *Physica A*, 333(0):497 – 504, 2004.

[146] A. A. Tsonis and K. L. Swanson. Topology and predictability of El Nino and La Nina networks. *Physical Review Letters*, 100(22):art.num.228502, 2008.

[147] A. A. Tsonis, K. L. Swanson, and P. J. Roebber. What do networks have to do with climate? *Bulletin of the American Meteorological Society*, 87(5):585–595, 2006.

[148] A. A. Tsonis, G. Wang, K. L. Swanson, F. A. Rodrigues, and L. F. Costa. Community structure and dynamics in climate networks. *Climate Dynamics*, 37(5-6):933–940, 2011.

[149] N. Tzourio-Mazoyer, B. Landeau, D. Papathanassiou, F. Crivello, O. Etard, N. Delcroix, B. Mazoyer, and M. Joliot. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage*, 15(1):273–289, 2002.

[150] H. VanLoon and J. C. Rogers. Seesaw in winter temperatures between greenland and northern Europe .1. General description. *Monthly Weather Review*, 106(3):296–310, 1978.

[151] J. M. Wallace and D. S. Gutzler. Teleconnections in the geopotential height field during the northern hemishere winter. *Monthly Weather Review*, 109(4):784–812, 1981.

[152] H. Wang, G. Yan, and Y. Xiao. Symmetry in world trade network. *Journal of System Science & Complexity*, 22(2):280–290, 2009.

[153] S. Wasserman and K. Faust. *Social Network Analysis*. Cambridge University Press, 1994.

[154] D. J. Watts and S. H. Strogatz. Collective dynamics of 'small-world' networks. *Nature*, 393(6684):440–442, 1998.

[155] N. Wiener. The theory of prediction. In *Modern Mathermatics for Engineers*, pages 165–190. McGraw-Hill: New York, USA, 1956.

[156] M. G. Wyatt, S. Kravtsov, and A. A. Tsonis. Atlantic Multidecadal Oscillation and Northern Hemisphere's climate variability. *Climate Dynamics*, 38(5-6):929–949, 2012.

[157] Y. Xiao, M. Xiong, W. Wang, and H. Wang. Emergence of symmetry in complex networks. *Physical Review E*, 77(6, 2):art.num.066108, 2008.

[158] Q. Xuan and T.-J. Wu. Node matching between complex networks. *Physical Review E*, 80(2):art.num.026103, 2009.

# List of Abbreviations

**Classes and properties of relational structures**

**AP** Amalgamation property, page 8

**EP** Extension property, page 9

$\mathcal{F}$ class of structures used to define class of forbidden structures, page 20

$\mathrm{Forb_h}(\mathcal{F})$ class of structures without homomorphic image of any $F \in \mathcal{F}$, page 20

$\mathrm{Forb}(\mathcal{F})$ class of structures omitting all structures from $\mathcal{F}$, page 20

$\mathcal{G}$ Gardiner graphs, page 17

**HH** Class of graphs extending every local homomorphism to global one, page 21

**HP** Hereditary property, page 8

**JEP** Joint embedding property, page 8

**HAP** homo-amalgamation property, page 24

**MAP** Mono-amalgamation property, page 23

**MH** Class of graphs extending every local monomorphism to global one, page 21

**MM** Class of graphs extending every local monomorphism to global one, page 21

**Examples of structures**

$P$ Universal ultrahomogeneous partial order, page 16

$\mathbb{Q}$ Rational numbers, page 7

**Graphs**

$\mathcal{C}$ set of all nonisomorphic cores of digraphs, page 15

$\overline{G}$ complement of graph $G$, page 12

$C_n$ cycle of size $n$, page 13

$D$ distance matrix, page 13

$d_{i,j}$ distance between vertices $v_i$ and $v_j$, page 13

$deg^+_G(v)$ out-degree of vertex $v$ in graph $G$, page 14

$deg^-_G(v)$ in-degree of vertex $v$ in graph $G$, page 14

$deg_G(v)$ degree of vertex $v$ in graph $G$, page 13

$diam(G)$ diameter of graph, page 13

$G, H$ digraphs, page 14

$\epsilon_G(v)$ eccentricity of graph, page 13

$\{v_i, v_j\}$ undirected edge of graph between vertices $v_i$ and $v_j$, page 3

$E, E(G)$ edge set of actual graph and edge set of graph $G$, page 3

$e_\ell, v_iv_j$ edges of graphs, either specific edge or using vertices concatenation, page 3

$G, H$ graph, page 3

$G \leq H$ there exists homomorphism between digraphs $G$ and $H$, page 15

$J_{n,k}$ Johnson graph, page 90

$K_n$ complete graph on $n$ vertices, page 12

$K_{m,n}$ complete bipartite graph, page 13

$K_{n_1,n_2,\ldots,n_m}$ complete multipartite graph, page 13

$KG_{n,k}$ Kneser graph, page 90

$L(G)$ line graph, page 14

$\Gamma(v)$ neighborhoodof vertex $v \in V(G)$, page 13

$\Gamma^+(v)$ out-neighborhood of vertex $v \in V(G)$, page 14

$\Gamma^-(v)$ in-neighborhood of vertex $v \in V(G)$, page 14

$\hom(F, G)$ number of homomorphism from $F$ to $G$, page 55

$P_n$ path of length $n$, page 13

$R$ Countable random graph, page 6

$G^\circ$ reflexive digraph, page 14

$T$ (graph) tree, page 13

$t(F, G)$ homomorphism density, page 55

$T_k$ transitive tournament, page 15

$V, V(G)$ vertex set of actual graph or vertex set of graph $G$, page 3

$v_i, u_j$ vertices of graphs, page 3

**$L$-colored graphs**

$\chi'$ vertex coloring function of $L$-colored graphs, page 15

$\chi''$ edge coloring function of $L$-colored graphs, page 15

$r, b$ colors used for bilocolored graphs, page 16

**Mathematical operations and symbols**

$\mathrm{Aut}(\mathbf{A})$ automorphism group of relational structure $\mathbf{A}$, page 4

$\mathrm{Inv}(\mathbf{A})$ union of $k$-ary invariants, i.e. $\mathrm{Inv}(\mathbf{A}) = \bigcup_{k\geq 1} \mathrm{Inv}_k(\mathbf{A})$, page 84

$\mathrm{Inv}_k(\mathbf{A})$ set of all $k$-ary invariants of $\mathrm{Aut}(\mathbf{A})$, page 84

$\mathrm{Inv}_{\leq k}(\mathbf{A})$ $\mathrm{Inv}_{\leq k}(\mathbf{A}) = \bigcup_{1\leq k'\leq k} \mathrm{Inv}_{k'}(\mathbf{A})$, page 84

$f = \binom{x_1\ x_2\ \ldots\ x_n}{y_1\ y_2\ \ldots\ y_n}$ mapping $f$ defined as $f(x_i) = y_i$, page 4

$\binom{X}{k}$ for $X$ being set, all subsets of $X$ with cardinality equal $k$, page 4

**Complex networks**

$r$ assortative coefficient, page 37

$C_b(i)$ betweenness centrality for vertex $i$, page 38

$\rho(X, Y)$ Pearson's correlation coefficient, page 40

$E$ efficiency, page 37

EEG electroencephalography, page 38

ENSO El Niño-Southern Oscillation, page 46

fMRI functional magnetic resonance imaging, page 38

$G(n, M)$ Erdős and Rényi random graph with $n$ vertices and $M$ edges, page 32

$G(n, p)$ Erdős and Rényi random graph with $n$ vertices and probability $p$, page 32

$I(X, Y)$ mutual information, page 40

$C$ clustering coefficient, page 36

$c_i$ local clustering coefficient, page 36

$L$ characteristic path length, page 36

PCA Principal Component Analysis, page 47

PET positron emission tomography, page 38

$\gamma$ small-world coefficient for clustering, page 36

$\lambda$ small-world coefficient for characteristic path, page 36

$\sigma$ small-world index, page 37

$W$ weighted connectivity matrix, page 39

**Probability theory**

$\mathbb{E}$ Expected value, page 40

**Relational structures**

$\mathrm{Age}\,(A)$ age of relational structure, page 8

$\mathbf{A}, \mathbf{B}, \mathbf{C}, \ldots$ relational structure, page 3

$A, B, C, \ldots$ domains of relational structures $\mathbf{A}, \mathbf{B}, \mathbf{C}, \ldots$, page 3

$\delta_i$ arity of $i^{th}$ relation of relational structure $\mathbf{A}$, page 3

$\mathbf{X} = (\mathbf{A}, X^1_{\mathbf{X}}, X^2_{\mathbf{X}}, \ldots, X^N_{\mathbf{X}})$ structure $X$ defined as extension of structure $A$ by relations $X^i_{\mathbf{X}}$ for $i = 1, 2, \ldots, N$, page 84

$a_i, x_i$ elements relational structure domain, page 3

$\varphi, \phi$ formulas or atomic formulas, page 19

$G_{\mathbf{A}}$ Gaifman graph of structure $\mathbf{A}$, page 92

$\mathrm{lc}(\mathbf{A})$ lift complexity of structure $\mathbf{A}$, page 84

$L$ signature of relational structure, page 3

$L_{\infty,0}$ first order language, page 19

$L_{\infty,0}$ quantifier-free language, page 19

$L_{\infty,\omega}$ language joining arbitrary many formulas and having finitely many quantifiers, page 19

$\mathbf{A} \models \varphi$ $\mathbf{A}$ is model of $\varphi$, page 19

$\overline{a}, \overline{x}$ $n$-tuple of elements, page 3

$\mathcal{A}$ rooted structure, page 94

$\mathrm{rc}(\mathbf{A})$ relational complexity of structure $\mathbf{A}$, page 84

$\mathrm{Rel}(L)$ class of all (countable) relational structures with signature $L$, page 3

$\mathbf{A} \mid_{L_1}$ $L_1$-reduct of $\mathbf{A}$, page 18

$\vec{R}$ root of rooted structure, page 94

$L(\mathbf{A})$ canonical lift, page 94

$R_{\mathbf{A}}$ tuple $(R^i_{\mathbf{A}} : i \in I)$ of all relations of structure $\mathbf{A}$, page 3

$R^i_{\mathbf{A}}$ $i^{th}$ relation of relational structure $\mathbf{A}$, page 3

$T, Th(\mathcal{K})$ theory, theory of class $\mathcal{K}$, page 19

$t_i, s_i$ terms of relational language, page 18

$U(n, \alpha, \beta)$ Uniform $L$-colored graph, page 66

# Glossary