SJÄLVSTÄNDIGA ARBETEN I MATEMATIK

MATEMATISKA INSTITUTIONEN, STOCKHOLMS UNIVERSITET

The Structure of Primitive Edge-Colored Graphs: A Galled Tree Perspective

av

Anna Lindeberg

2023 - M7

MATEMATISKA INSTITUTIONEN, STOCKHOLMS UNIVERSITET, 106 91 STOCKHOLM

The Structure of Primitive Edge-Colored Graphs: A Galled Tree Perspective

Anna Lindeberg

Självständigt arbete i matematik 30 högskolepoäng, avancerad nivå

Handledare: Marc Hellmuth

2023

Sammanfattning

Ett gallat träd är en rotad, riktad och acyklisk graf där inga två 'oriktade' cykler har en gemensam kant. En komplett kantfärgad graf Σ förklaras av ett gallat träd (N, t) om varje nod i Σ har ett motsvarande löv i N, och funktionen t förser varje nod i N med en färg-etikett på så sätt att den minsta gemensamma förfädern av varje par av löv x och y antar samma färg som kanten $\{x, y\}$ i Σ. Vi karakteriserar först vilka kompletta kantfärgade grafer av en viss typ (så kallade primitiva kantfärgade grafer) som kan förklaras av ett gallat träd. Efter detta undersöker vi när det finns ett unikt gallat träd som förklarar en viss primitiv kantfärgad graf.

Abstract

A galled tree is a rooted, directed, and acyclic graph such that no two 'undirected' cycles in it share an edge. A complete edge-colored graph Σ is explained by a labeled galled tree (N, t) if the vertices of Σ are the leaves of N, and t assigns a label to each vertex of N in such a way that the label of the least common ancestor of any two leaves x and y equals the edge-color of the edge $\{x, y\}$ in Σ. In this thesis we characterize which complete edge-colored graphs of a particular type (so-called primitive edge-colored graphs) can be explained by a galled tree. Furthermore, we investigate when such a galled tree is uniquely determined.

Contents

1 Introduction

Over the years, mathematicians and computer scientists have both searched for and found a wide variety of graph classes united by the property that some generally 'hard' problems are possible to solve efficiently on members of those classes. One example of such a class of graphs are the so-called *cographs*. A graph is a cograph if and only if there exists a rooted tree, eloquently known as a cotree, whose set of leaves coincides with the vertices of its cograph. Additionally, the cotree's inner vertices are labeled with 0:s and 1:s such that the least common ancestor of two leaves x and y is labeled 1 if and only if $\{x, y\}$ is an edge of the cograph. We give an example of a cograph and its cotree in Figure 1. In fact, the existence of the cotree is what makes cographs interesting from an algorithmic point of view; the cotree captures exactly the same information as its corresponding cograph, so one may run algorithms on the cotree to get solutions for the corresponding cograph. Examples of NP-hard problems that are polynomial time-solvable on cographs include finding the size of a maximum clique and deciding Hamiltonicity. Furthermore, whether two cographs are isomorphic or not can be decided in polynomial time [4].

The closest object resembling a cotree which is uniquely determined for every possible graph is something called the *modular decomposition tree* (MDT) of a graph [15]. Like cotrees, the MDT of a graph G is a rooted tree whose leaves coincides with the vertices of G, and the vertices of the MDT are labeled. However, for non-cographs the labels 0 and 1 does not suffice, and instead a third label Prime is introduced. The MDT of a graph G often captures some structure of G, in the sense that if the label of the least common ancestor of two leaves x and y equals 1 (respectively 0), then $\{x, y\}$ is (respectively is not) an edge of G. However, if the least common ancestor of x and y is labeled Prime then nothing can be derived about the (non-)edge $\{x, y\}$ in G. The existence of Prime-labeled vertices can thus be thought of as a roadblock in efficiently running algorithms on the MDT of a graph. For cographs, the cotree coincides with the MDT which is thus devoid of Prime-labels. We say that the MDT of a graph G without Prime-labels (that is, G is a cograph) *explains* G.

When it comes to how much of a graph G can be recovered from its MDT the other, so to speak, 'extreme' than cographs are the *primitive* graphs. A graph is primitive if it has only one or two vertices, or if its MDT consists of a Prime-labeled root that has only children that are leafs. Hence virtually no information about the underlying graph can be retrieved from the MDT of a primitive graph. The smallest non-trivial example of a primitive graph is a path graph P_4 with four vertices, depicted in Figure 2.

Although the long history of study of modular decomposition, the surprisingly simple idea of replacing Prime-labeled vertices by some other well-understood type of graph appeared only recently [3]. Perhaps the most 'tree-

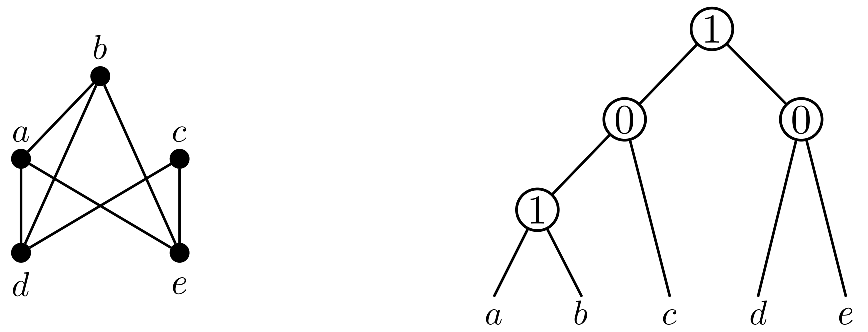


Figure 1: A cograph (left) with its cotree (right). By way of example, note that the least common ancestor of a and d is labeled 1, and the edge $\{a, d\}$ indeed appear in the graph.

like' object that is not a tree is what is known as a *galled tree*. We introduce these objects formally in the next section, but intuitively speaking a galled tree is a rooted tree where cycles (also known as galls) are allowed, as long as no two cycles share an edge.

In the recent work [19] Hellmuth and Scholz generalized cographs to what they call GaTEx graphs. A GaTEx (Galled Tree Explainable) graph is a graph that can be explained by a labeled galled tree, analogously to how cographs can be explained by labeled trees. For example, the non-cograph P_4 is a GaTEx graph, as it is explained by the galled tree in Figure 2. The work in [19] is extensive; to only mention a few results it first characterizes graphs that are both primitive and GaTEx, and then use a carefully defined type of quotient graphs to extend this characterization to all GaTEx graphs. In fact, the same approach as the Prime-vertex replacement introduced in [3] is used for the latter. These major results are then complemented with results of uniqueness-type, i.e. which GaTEx graphs are explained by a unique (up to isomorphism) galled tree. Moreover, algorithms that both recognize GaTEx graphs and construct a labeled galled tree that explain a particular GaTEx graph are given, and these algorithms are implementable in linear (i.e. polynomial) time.

The work in [21] complements [19] by providing a characterization of GaTEx graphs in terms of so-called forbidden subgraphs, and uses this to establish connections between GaTEx graphs and many other well-known graph classes (e.g. perfect graphs and comparability graphs). In the third recent paper [20] of Hellmuth and Scholz, the authors further motivates their previous work by providing linear-time algorithms which solves problems that in general are NP-hard, as long as the input is restricted to a GaTEx graph. The common approach they use for computing the size of a maximum clique, the size of a maximum independent set respectively the number of colors in an optimal vertex-coloring of a given GaTEx graph G is that the

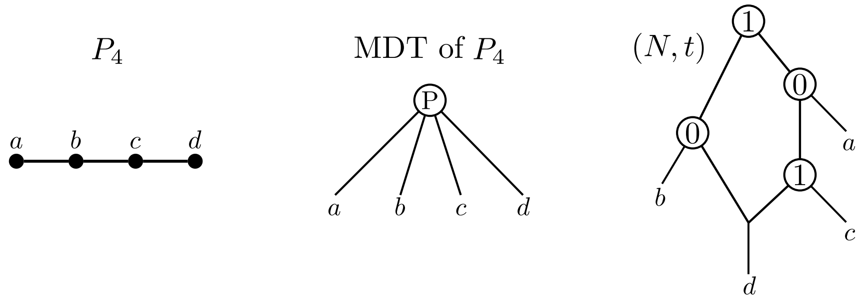


Figure 2: The path graph P_4 on four vertices (left) is primitive, as its MDT (middle) has a PRIME-labeled root and four leaves only. It is, however, explained by the galled tree (N, t) (right) and is thus a GATEX graph.

algorithm in question is run on a labeled galled tree that explains G, rather than on G itself. In other words it is, so to say, easy to solve hard problems on GATEX graphs for much the same reasons as it is easy to solve hard problems on cographs.

In this contribution, we wish to generalize parts of the results of [19] to another combinatorial object than undirected graphs, namely, to complete edge-colored graphs. These are, as the name indicates, complete graphs where each edge is assigned an edge-color. Thus the binary property of $\{x, y\}$ either being an edge of a graph G, or not, is exchanged to the more general situation where each edge has one of many possible edge-colors. In this context, a complete edge-colored graph G is explained by a labeled galled tree (N, t), where N is the galled tree and t its labeling, if the label $t(\text{lca}(x, y))$ equals the edge-color of the edge $\{x, y\}$, for each edge $\{x, y\}$ of G. That is, instead of t only assigning 0/1-labels to the vertices of N, it assigns colors to the vertices of N. The concept of modular decomposition trees also generalizes to complete edge-colored graphs, in a similar fashion. In particular, no primitive complete edge-colored graph on at least three vertices can be recovered from its MDT.

We do not pretend to fully generalize every result in [19], but we will provide the first stepping-stones towards such a possibility. To be more precise, we will first investigate complete edge-colored graphs that are explained by a particularly ‘simple’ type of galled tree (called *elementary*) that, roughly speaking, has a single cycle as a ‘backbone’ with separated leaves – see Figure 3 for an example. As it turns out, complete edge-colored graphs can be explained by an elementary galled tree (with some additional conditions of technical nature) if and only if it is primitive yet explained by *some* galled tree. To facilitate the proof of this characterization, we will introduce the class of so-called *polar-cats*. We state the formal definition of polar-cats in Section 3, but the rough idea is that a complete edge-colored graph is a polar-cat if it can be split into two subgraphs that intersect only

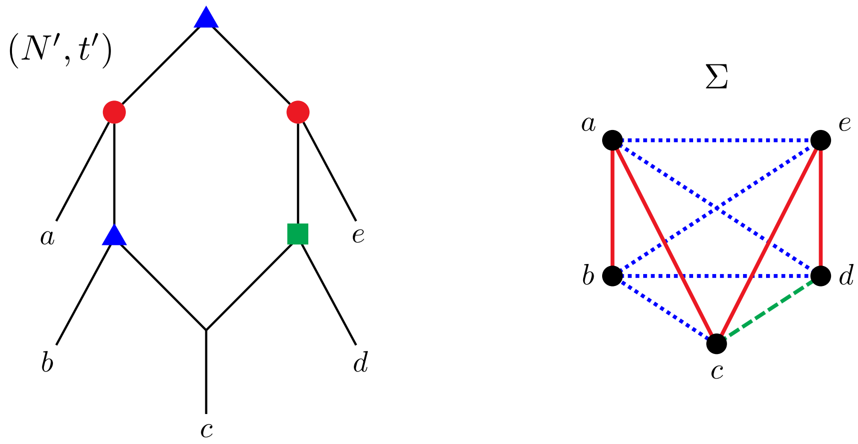


Figure 3: We give an elementary labeled galled tree (N', t') (left) that explains the complete edge-colored graph Σ (right). For example, the least common ancestor of the leaves c and d is labeled ■, and the edge $\{c, d\}$ is indeed dashed green in Σ.

in one vertex, and each of those subgraphs have a very particular structure. This structure will allow us to prove that polar-cats are both explained by particular galled trees and primitive, thus acting as a sort of middle ground for the main theorem of Section 3.

Just like cographs are the class of undirected graphs that can be explained by a tree, the so-called *uniformly non-primitive* ($\mathfrak{unp}$) complete edge-colored graphs can be explained by a labeled tree. Equivalently, a complete edge-colored graph is $\mathfrak{unp}$ if and only if its MDT has no PRIME-labeled vertices. Modular decomposition trees and $\mathfrak{unp}$ edge-colored graphs have previously been studied extensively, although in the yet even more general context of complete edge-colored *directed* graphs (also known as 2-structures). Important previous work can be found in [7, 8, 9, 11, 22]. In particular, [22] connects the previously theoretical study of 2-structures to the field of computational biology, as relationships between different genes can be modeled with edge-colored directed graphs.

As previously mentioned, we state and prove a characterization of primitive complete edge-colored graphs that are explained by labeled galled trees in Section 3. After that, we devote Section 4 to a characterization of which primitive complete edge-colored graphs are explained by a *unique* galled tree. Section 2 provides the relevant background on graphs, complete edge-colored graphs, galled trees and modular decomposition trees. The reader familiar with the works in [19] may safely read only Section 2.2 for the formal definitions involving complete edge-colored graphs, and refer back to the remainders of Section 2 when necessary. We conclude this thesis with an outlook in Section 5.

2 Preliminaries

In this section we formally introduce a substantial collection of definitions and earlier results needed later on. We begin by summarizing basic facts and notation related to directed and undirected graphs, followed by a discussion about so-called 2-structures and their connection to edge-colored graphs. We then formalize the important galled trees discussed in the introduction and establish a rather extensive set of definitions related to this type of directed network. Lastly, we provide an overview on the topic of modular decomposition and its connection to galled trees.

Before anything else, let us establish some basics about sets. All sets considered in this thesis are assumed to be finite. For a nonempty set V, $\binom{V}{2}$ denotes the family of two-element subsets of V. A *partition of a set* X is a family of nonempty sets $\{X_1, \ldots, X_k\}$ such that $X = \cup_{i=1}^{k} X_i$ and such that $X_i \cap X_j = \emptyset$ for each $i \neq j$.

2.1 Graphs

The concepts in this section are well-known, but notation varies largely. For a more in-depth discussion about the topics discussed, refer to any textbook in graph theory, for example [6]. A *graph* $G = (V, E)$ is an ordered pair consisting of a nonempty set $V(G) := V$ of *vertices* and a set $E(G) := E$ of *edges*. $G = (V, E)$ is an *undirected graph* if $E \subseteq \binom{V}{2}$. If, instead, $E \subseteq (V \times V) \setminus \{(v, v) \,|\, v \in V\}$, then G is said to be a *directed graph*. We thus distinguish between graphs with *undirected edges* $\{x, y\}$ and graphs with *directed edges* (x, y), for distinct $x, y \in V$. In particular, the given definition ensures that all graphs we consider have no loops (i.e. edges such as $\{x, x\}$ or (x, x)) nor any multi-edges (i.e. multiple occurrences of the same edge).

The vertices x and y are called *endpoints* of the edge e, when $e = \{x, y\}$ or $e = (x, y)$, while x and y are *incident* to the edge e. If x and y are vertices connected by an (undirected or directed) edge, then x and y are said to be *adjacent* or *neighbors*. We will consider both undirected and directed graphs, but the latter only in the context of galled trees, see Section 2.3. Thus, when we refer to a graph only, we implicitly mean an undirected graph.

Example 2.1. An example of an (undirected) graph $G = (V, E)$ is given in (the left of) Figure 4. As customary, vertices are drawn as black circles and often labeled, while (undirected) edges are indicated with lines. In G, a and b are adjacent vertices, whereas a and e are not, since $\{a, b\} \in E$ while $\{a, e\} \notin E$.

Given two graphs $G = (V_G, E_G)$ and $H = (V_H, E_H)$ (either both directed or both undirected), an *isomorphism of* G *and* H is a bijective map $\varphi : V_G \to V_H$ such that $\{u, v\} \in E_G$ (resp. $(u, v) \in E_G$) if and only if $\{\varphi(u), \varphi(v)\} \in E_H$ (resp. $(\varphi(u), \varphi(v)) \in E_H$). If such a map exist, then G and H are *isomorphic*, in symbols stated as $G \simeq H$.

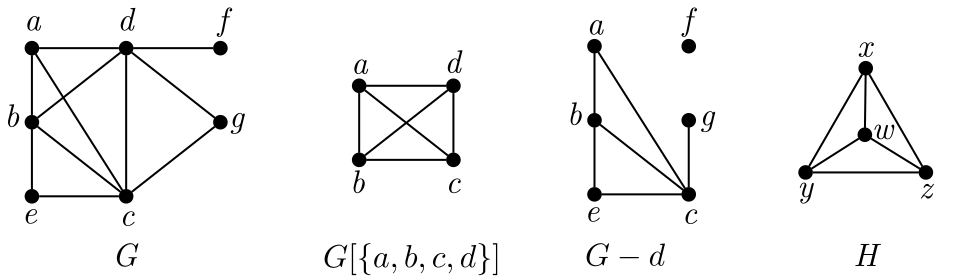


Figure 4: Example of a graph G (left), an induced subgraph $G[\{a,b,c,d\}]$ (left middle), the graph $G-d$ (right middle) and a graph H. Note that $G[\{a,b,c,d\}]$ is isomorphic to the (complete) graph H on four vertices.

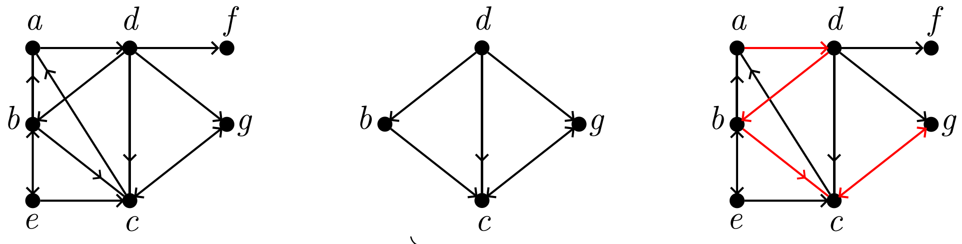
A *subgraph* of a (directed or undirected) graph $G=(V,E)$ is a graph $G'=(V',E')$ such that $V'\subseteq V$, $E'\subseteq E$ and such that for every $e\in E'$, both endpoints of e are elements of V'. Given a nonempty subset of vertices $X\subseteq V$ of the (undirected or directed) graph $G=(V,E)$, the *subgraph induced by* X is the (sub-)graph with vertex set X and edge set

$$E'=\{e\in E\,|\,\text{both endpoints of } e \text{ lies in } X.\}$$

This graph is denoted by $G[X]$. In particular, we put $G-v:=G[V\setminus\{v\}]$ whenever $V\setminus\{v\}\neq\emptyset$. A graph H is an *induced subgraph* of G if there exists some set X such that $H\simeq G[X]$. In that case, we also say that G *contains* H.

Example 2.2. Continuing Example 2.1, the two induced subgraphs $G':=G[\{a,b,c,d\}]$ and $G-d$ are given alongside G in Figure 4. Moreover, G' is isomorphic to the graph H depicted in (the right of) Figure 4. In fact, it is easy to verify that any injective map from $V(G')$ into $V(H)$ indeed defines an explicit isomorphism of these two graphs.

A graph $G=(V,E)$ is *complete* if $E=\binom{V}{2}$. For example, the graph H of Figure 4 is a complete graph on four vertices. For each positive integer n, an *undirected path* P_n is the undirected graph with vertices $\{v_1,v_2,\dots,v_n\}$ and edges $\{v_k,v_{k+1}\}$ for each $k=1,2,\dots,k-1$. Similarly, a *directed path* $\overrightarrow{P}_n$ is the directed graph with vertices $\{v_1,v_2,\dots,v_n\}$ and edges (v_k,v_{k+1}) for each $k=1,2,\dots,k-1$. Whether a path is directed or undirected is often understood from context, and they are at times denoted with its inner vertices, so that $P_n=v_1v_2\dots v_n$. The paths P_n and $\overrightarrow{P}_n$ are said to have *length* $n-1$, the vertices v_1 and v_n are called the *end vertices* of the path while the vertices v_k for $k\notin\{1,n\}$ are called *inner vertices*. In particular, $n=1$ means that the path in question consists of a single (end) vertex, so that it has length zero.



$\overrightarrow{G}$ $\quad$ $\overrightarrow{G}[\{b,c,d,g\}]$

Figure 5: Example of a directed graph $\overrightarrow{G}$ (left), its induced subgraph $\overrightarrow{G}[\{b,c,d,g\}]$ (middle) and a highlighted ag-path in $\overrightarrow{G}$ (right). Note that $\overrightarrow{G}[\{b,c,d,g\}]$ is both biconnected and acyclic, two properties not satisfied by $\overrightarrow{G}$.

Example 2.3. Once again considering the graph G in Figure 4, we see that the vertices b, d, e and f induce a path graph on four vertices, so G contains a P_4. In $G[\{b,d,e,f\}]$, e and f are end vertices, while b and d are inner vertices.

Given a directed or undirected graph G and two of its vertices x and y, there is a *path in G (from x to y)* or an *xy-path* for short, if there exist a subgraph (not necessarily induced) of G isomorphic to P_n (respectively $\overrightarrow{P}_n$, if G is directed) for some n so that x and y are its end vertices. We may also consider undirected paths in directed graphs as follows. For any directed graph $\overrightarrow{G} = (V,E)$, define the undirected graph $\overrightarrow{G}^u := (V,E^u)$ where $E^u := \{\{x,y\} \mid (x,y) \in E\}$. An *undirected xy-path in $\overrightarrow{G}$* is any xy-path in $\overrightarrow{G}^u$.

Example 2.4. Consider the directed graph $\overrightarrow{G} = (V,E)$ as given in Figure 5. The direction of the edges are indicated with arrows. As highlighted in red, there is, for example, a directed path of length four in $\overrightarrow{G}$ from a to g, via the vertices d, b and c (in that order). On the other hand, there is no directed path from f in $\overrightarrow{G}$, since there is no $x \in V$ such that $(f,x) \in E$. However, there is an undirected fx-path for each $x \in V$.

A (directed or undirected) graph G is *connected* if there exist an (undirected) xy-path between any pair of vertices x and y. If G has only one vertex, or if $G - v$ is connected for each vertex v of G, then G is said to be *biconnected*. A *biconnected component* $C = (V',E')$ of $G = (V,E)$ is a maximal biconnected subgraph. That is, C is biconnected while, for any $x \in V \setminus V'$, the subgraph $G[V' \cup \{x\}]$ is not biconnected. A biconnected component is *non-trivial* if it has at least three vertices, otherwise it is *trivial*.

Example 2.5. The graph $\overrightarrow{G}$ in Figure 5 has several biconnected components: two examples are $\overrightarrow{G}[\{d, f\}]$ and $\overrightarrow{G} - f$. The latter is non-trivial, while the former is trivial. The subgraph $\overrightarrow{G}[\{b, c, d, g\}]$ also depicted in Figure 5 is biconnected, but not a biconnected component, since for example $\overrightarrow{G}[\{b, c, d, g\} \cup \{a\}]$ is biconnected as well.

The *degree* of a vertex v of an undirected graph is the number of neighbors of v. Let $G = (V, E)$ be a directed graph. For each $v \in V$ we define

$$\mathrm{indeg}(v) := |\{u : (u, v) \in E\}| \quad \text{and} \quad \mathrm{outdeg}(v) := |\{u : (v, u) \in E\}|$$

as the *in-degree* respectively *out-degree* of v. Moreover, G is said to be *acyclic* if there exist no sequence of $k \geq 2$ distinct vertices $v_1, v_2, \ldots, v_k \in V$ such that

$$(v_1, v_2), (v_2, v_3), \ldots, (v_{k-1}, v_k), (v_k, v_1) \in E.$$

Example 2.6. In our reoccurring example $\overrightarrow{G} = (V, E)$ in Figure 5, we have $\mathrm{indeg}(a) = 2$ and $\mathrm{outdeg}(a) = 1$ while $\mathrm{indeg}(f) = 1$ and $\mathrm{outdeg}(f) = 0$. $\overrightarrow{G}$ is not acyclic, since (among others) the sequence a, d and b satisfies $(a, d), (d, b), (b, a) \in E$. The subgraph $\overrightarrow{G}[\{b, c, d, g\}]$ is, on the other hand, indeed acyclic.

If $G = (V, E)$ and $H = (V', E')$ are graphs such that $V \cap V' = \emptyset$, then the *disjoint union of* G *and* H is the graph $G \cup H := (V \cup V', E \cup E')$. The *complement* $\overline{G}$ of the undirected graph $G = (V, E)$ is the graph with vertex set $V(\overline{G}) := V$ and edge set $E(\overline{G}) := \binom{V}{2} \setminus E$. In other words, the edges of $\overline{G}$ are precisely the non-edges of G.

2.2 Complete edge-colored graphs

One way to think about undirected graphs is that they represent a type of binary relationship between its vertices: two vertices are somehow related (depending on the context the graph is used in) if they share an edge, and not related if they do not share an edge. A natural generalization of this is to allow more than one type of edge and one type of non-edge; we can think of it as assigning a label or color to each pair of vertices. The formal definition is the following. We define a *(complete) edge-colored graph* as an ordered pair $\Sigma = (V, \sigma)$, where V is a (nonempty) set of *vertices* and σ is a map from $\binom{V}{2}$. Elements of $\binom{V}{2}$ are, unsurprisingly, called *edges*. If the codomain of σ needs to be referred to specifically, it is called the *edge-colors* of Σ. To avoid redundancy, we assume that σ is surjective. That means that when we state e.g. that Σ is an edge-colored graph with three edge-colors, then we assume that σ is a surjection into a codomain of cardinality three. For brevity, we often write $\sigma(xy)$ rather than $\sigma(\{x, y\})$.

$\Sigma = (V, \sigma)$ $\qquad\qquad$ $\Sigma' = (V, \sigma')$

$\sigma : \binom{V}{2} \to \{\text{●}, \text{▲}\}$ $\qquad\qquad$ $\sigma' : \binom{V}{2} \to \{\text{●}, \text{■}, \text{▲}\}$

Figure 6: Examples of two edge-colored graphs Σ and Σ'.

To avoid being repetitive, we write only 'edge-colored graph' instead of 'complete edge-colored' graph. Edge-colored graphs can be visualized much like graphs, as long as the edges between vertices are depicted with distinct colors. Examples of two edge-colored graphs Σ and Σ' are given in Figure 6. Σ has two edge-colors, whereas Σ'' has three.

We now introduce a handful of definitions for edge-colored graphs. Given an edge-colored graph $\Sigma = (V, \sigma)$ and a nonempty set $X \subseteq V$, we define the *induced subgraph* $\Sigma[X]$ *of* Σ to be the edge-colored graph $(X, \sigma|_X)$, where $\sigma|_X$ is the restriction of σ to the domain $\binom{X}{2}$. All subgraphs of edge-colored graphs are assumed to be induced subgraphs, deviating from the case of graphs without edge-colors. As for graphs, $\Sigma - v$ denotes the subgraph $\Sigma[V \setminus \{v\}]$.

We extend the definition of isomorphy of edge-colored graphs by adapting the terminology of [8]. We say that two edge-colored graphs $\Sigma = (V, \sigma)$ and $\Sigma' = (V', \sigma')$ are *isomorphic*, i.e. $\Sigma \simeq \Sigma'$, if there exist a bijective map $\varphi : \binom{V}{2} \to \binom{V'}{2}$ such that

$$\sigma(e) = \sigma(f) \iff \sigma'(\varphi(e)) = \sigma'(\varphi(f))$$

for all $e, f \in \binom{V}{2}$.

Given two vertex disjoint edge-colored graphs $\Sigma = (V, \sigma)$ and $\Sigma' = (V', \sigma')$ we can form a new edge-colored graph $\Sigma \otimes_k \Sigma'$ called the *k-colored join* of Σ and Σ'. It is the edge-colored graph with vertex set $V \cup V'$, and its coloring map $\sigma_{\Sigma \otimes_k \Sigma'}$ is defined by putting

$$\sigma_{\Sigma \otimes_k \Sigma'}(xy) := \sigma_{\Sigma \otimes_k \Sigma'}(\{x, y\}) := \begin{cases} \sigma(xy) & \text{if } x, y \in V \\ \sigma'(xy) & \text{if } x, y \in V' \\ k & \text{otherwise.} \end{cases}$$

for each pair of distinct elements $x, y \in V \cup V'$.

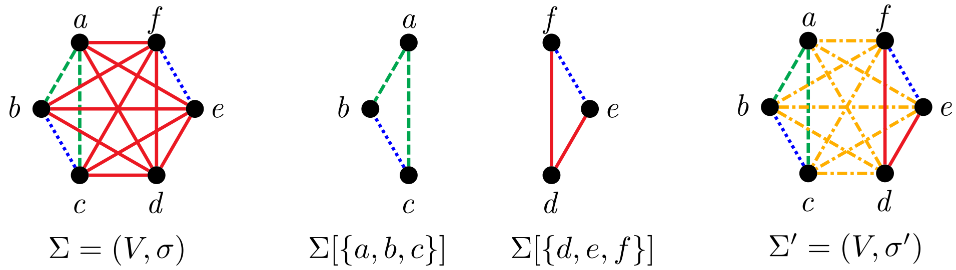


Figure 7: The graph Σ (left) is the ${\color{red}\bullet}$-colored join of its subgraphs $\Sigma[\{a, b, c\}]$ and $\Sigma[\{d, e, f\}]$ (middle), respectively. Similarly, Σ' (right) is the ${\color{orange}\blacktriangledown}$-colored join of $\Sigma'[\{a, b, c\}] = \Sigma[\{a, b, c\}]$ and $\Sigma'[\{d, e, f\}] = \Sigma[\{d, e, f\}]$.

Example 2.7. An edge-colored graph Σ and its two subgraphs $\Sigma[\{a, b, c\}]$ respectively $\Sigma[\{d, e, f\}]$ are given in Figure 7. It is easy to verify that

$$\begin{aligned}\Sigma &= \Sigma[\{a, b, c\}] \otimes_{\color{red}\bullet} \Sigma[\{d, e, f\}] \\ &= \Sigma[\{a, b, c, d\}] \otimes_{\color{red}\bullet} \Sigma[\{e, f\}] \\ &= (\Sigma - d) \otimes_{\color{red}\bullet} \Sigma[\{d\}],\end{aligned}$$

where ${\color{red}\bullet}$ is the color of the solid red edges in the figure. Figure 7 also depicts Σ', which satisfies

$$\Sigma' = \Sigma'[\{a, b, c\}] \otimes_{\color{orange}\blacktriangledown} \Sigma'[\{d, e, f\}].$$

It is sometimes of interest to consider only the edges of a single color of an edge-colored graph $\Sigma = (V, \sigma)$. For any edge-color k of Σ, the *k-colored monochromatic subgraph of* Σ, denoted $\Sigma|_k$, is the graph (V, E) where

$$E = \left\{ e \in \binom{V}{2} \,\middle|\, \sigma(e) = k \right\}.$$

Example 2.8. Continuing with the edge-colored graph Σ from Example 2.7 and Figure 7, its three monochromatic subgraphs are given in Figure 8. Note that $\Sigma|_{\color{red}\bullet}$, $\Sigma|_{\color{blue}\blacktriangle}$ and $\Sigma|_{\color{green}\blacksquare}$ are graphs without edge-colors, all with the same vertex set as Σ.

2.2.1 Related concepts

We now introduce the even more generalized notion of 2-structures, and connect edge-colored graphs to (undirected) graphs. The reader should be aware that the definitions in this section will only occasionally appear in the remainder of this thesis. They are nevertheless important to discuss, as they connect the edge-colored graphs to other combinatorial structures and thus a wider set of related results.

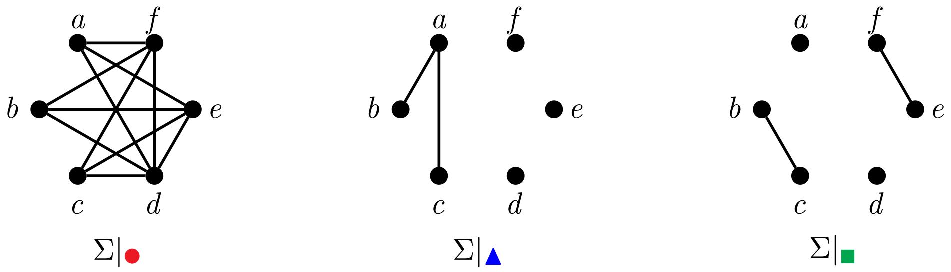


Figure 8: The three possible monochromatic subgraphs of the edge-colored graph Σ of Figure 7.

The theory of so-called 2-structures was introduced in [8, 9] as a combinatorial object that can both be seen as a generalization of (directed) graphs and a restriction of, for example, posets. Given a nonempty set V, we let $V^2_{\text{irr}} = (V \times V) \setminus \{(v,v) \mid v \in V\}$. In the context of 2-structures, the elements of V^2_{irr} are called *arcs*. A *2-structure* is an ordered pair $g = (V, \sim)$, where V is a nonempty set consisting of the *vertices* of g, and $\sim$ is an equivalence relation on V^2_{irr}. Since equivalence relations induce a partition of its underlying set (see e.g. [16, p.28]), an equivalent definition is that a 2-structure is a set of vertices V together with a partition $\mathcal{P}$ of V^2_{irr}, written as $g = (V, \mathcal{P})$. A third (in essence) equivalent notion of 2-structures uses the notion of a *labeling function* $\varphi : V^2_{\text{irr}} \to \Upsilon$, which assigns a label $\varphi((x,y)) \in \Upsilon$ to each arc $(x,y) \in V^2_{\text{irr}}$. This implicitly partitions V^2_{irr} in classes of arcs all labeled with the same element of Υ, so that the 2-structure $g = (V, \mathcal{P})$ can be identified with the triple $g = (V, \Upsilon, \varphi)$, as has been done in for example [22]. In [8, 9], these objects are distinguished with the name *labeled 2-structure*. The three notions all have their respective advantages, and their usefulness depends on the context. In particular, 2-structures are not dependent on the elements of the label set Υ; the importance lies in establishing the (non-)relationship between its arcs.

A 2-structure $g = (V, \sim)$ is said to be *symmetric* if $(x,y) \sim (y,x)$ for all arcs (x,y) in V^2_{irr}. Intuitively speaking, this makes sure that the "directed" arcs (x,y) and (y,x) are always grouped together into an "undirected" *edge* $\{x,y\}$. In particular, any labeled symmetric 2-structure $g = (V, \Upsilon, \varphi)$ can be reinterpreted without loss of information as the complete edge-colored graph $\Sigma = (V', \sigma)$, where $V' = V$ and $\sigma(\{x,y\}) = \varphi((x,y)) = \varphi((y,x))$ for all distinct vertices x and y, and *vice versa*. Although it is not true to state that the definition of symmetric 2-structures is equivalent to the definition of complete edge-colored graphs, it is at least evident that there is a one-to-one correspondence between these two families of objects. In particular, any result on (symmetric) 2-structures is also true for complete edge-colored graphs.

The study of (vertex or edge) colored graphs is a prominent field within graph theory. One often studies undirected graphs $G = (V, E)$ *equipped with an edge-coloring* $\sigma : E \to \Upsilon$, where Υ is a set of colors. The notion of complete edge-colored graphs we use can clearly be seen as a special case of this; a complete edge-colored graph $\Sigma = (V, \sigma)$ is the same as considering the complete graph $G = (V, \binom{V}{2})$ equipped with the edge-coloring σ. Perhaps more surprisingly, the converse is also true, in the following sense. If $G = (V, E)$ is a graph (not necessarily complete) equipped with an edge-coloring σ, then we may construct a complete edge-colored graph $\Sigma_G = (V', \sigma')$ that 'encodes' the same information as (G, σ) by putting $V' := V$ and

$$\sigma'(e) := \begin{cases} \sigma(e) & \text{if } e \in E \\ \circ & \text{otherwise} \end{cases}$$

for each $e \in \binom{V}{2}$. Here $\circ$ is assumed to be any symbol not already in the codomain of σ. The difference between complete edge-colored graphs and graphs equipped with an edge-coloring is thus smaller than one might believe at a first glance; in graphs equipped with an edge-coloring we, so to say, keep track of edge-colors and non-edges, while in complete edge-colored graphs we equate being a non-edge with having some particular color not appearing elsewhere.

Another important point is that the results we will develop in this thesis may be transferred to the case of undirected graphs. The easiest way to see this is to note that we may always equip a graph $G = (V, E)$ with a constant edge-coloring σ, i.e. one that satisfies $\sigma(e) = \sigma(f)$ for all $e, f \in E$. With that edge-coloring, the discussion in the previous paragraph directly applies, yet σ does not distinguish between edges any more than G itself does.

Example 2.9. The graph $G = (V, E)$ in Figure 9 can be identified with the edge-colored graph $\Sigma = (V, \sigma)$ of Figure 6 by constructing σ such that

$$\sigma(e) = \begin{cases} \bullet & \text{if } e \in E \\ \blacktriangle & \text{otherwise} \end{cases}$$

for each $e \in \binom{V}{2}$. Visually, this means edges of G correspond to solid red edges of Σ, while non-edges of G correspond to dotted blue edges of Σ.

If G is equipped with an edge-coloring map such as σ'', depicted in Figure 9 with solid red respectively dashed green edges, we can reinterpret the pair (G, σ'') as the edge-colored graph $\Sigma' = (V, \sigma')$ of Figure 6. In that case, σ satisfies

$$\sigma(e) = \begin{cases} \sigma''(e) & \text{if } e \in E \\ \blacktriangle & \text{otherwise} \end{cases}$$

for each $e \in \binom{V}{2}$.

Figure 9: Examples of a graph G with no edge-colors (left) that can be identified with the edge-colored graph Σ of Figure 6. If G is equipped with an edge-coloring map σ'' (right) then its reinterpretation as an edge-colored graph is Σ' of Figure 6. See text and, in particular, Example 2.9 for details.

It is worth reiterating that the definition of edge-colored graphs used in this thesis thus includes both general (undirected) graphs and non-complete graphs equipped with an edge-coloring. The definition is also essentially the same as that of symmetric (labeled) 2-structures, and thus less general than arbitrary 2-structures. The choice of the terminology is motivated by the close connections to the work in [19, 21], where graphs (without edge-colors) that can be explained by galled trees are characterized.

2.3 Galled trees

Many different types of phylogenetic trees and networks have been considered, adapting the definitions to the context of the research in question. See [24] for an overview. We will follow the definitions of [19] quite closely, but adapt some of the terminology to that of [21]. Given a nonempty set X, a *galled tree N (on X)* is a directed acyclic graph (V, E) such that either

(N0) $V = X = \{x\}$ and, therefore, $E = \emptyset$

or such that N satisfies the following properties:

(N1) N has a unique *root* denoted ρ_N such that $\operatorname{indeg}(\rho_N) = 0$ and $\operatorname{outdeg}(\rho_N) \geq 2$, and

(N2) $x \in X$ if and only if $\operatorname{indeg}(x) = 1$ and $\operatorname{outdeg}(x) = 0$, and

(N3) every $v \in V \setminus X$ with $v \neq \rho_N$ is a

  (a) *tree vertex* i.e. satisfies $\operatorname{indeg}(v) = 1$ and $\operatorname{outdeg}(v) \geq 2$, or a
  (b) *hybrid vertex* i.e. satisfies $\operatorname{indeg}(v) = 2$ and $\operatorname{outdeg}(v) \geq 1$

  and, lastly,

(N4) each biconnected component C of N has at most one hybrid vertex η such that $x \in V(C)$ for some vertex x with $(x, \eta) \in E$.

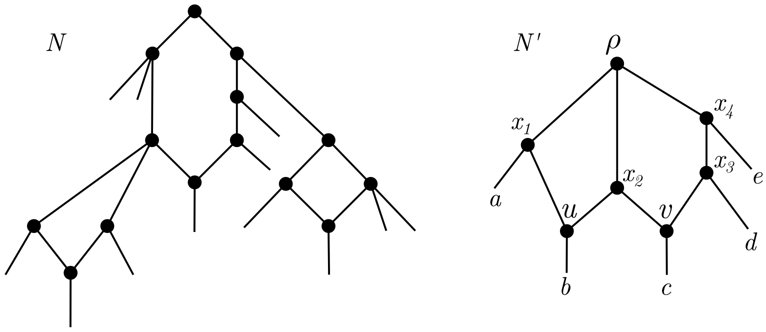


Figure 10: An example of a galled tree N (left) and a phylogenetic network N' (right) that is not a galled tree, since it fails condition (N4).

Members of the set X are called the *leaves* of N, sometimes also denoted $\mathcal{L}(N)$. The set $V^0(N) := V \setminus X$ contains the *inner vertices* of N. As noted in [19] condition (N4) implies that the vertices of each non-trivial biconnected component C of a galled tree can be divided into two directed paths P^1 and P^2 which intersect only at their endpoints. Such a biconnected component is called a *cycle*. The paths P^1 and P^2 are called the *sides* of the cycle C. By definition, a cycle C will have a unique vertex ρ_C with in-degree zero called its *root*, and a unique vertex η_C with out-degree zero called its *hybrid vertex*. Cycles in galled trees are also known under the names *galls* [13] and *blocks* [14]. Note also that (N4) is a property known as N being *level-1* [19].

Example 2.10. We give an example of a galled tree N to the left in Figure 10. To avoid cluttering, the direction of the edges are not indicated with arrows, instead the direction of each edge is always assumed to be from top to bottom. Inner vertices are indicated with black dots, while leaves appear at the end of edges without black dots. In particular, N has twelve leafs, 14 inner vertices and three cycles.

To the right in Figure 10 an acyclic directed graph $N' = (V, E)$ is given. N' has five leafs, namely $\mathcal{L}(N) = \{a, b, c, d, e\}$ and its inner vertices $V^0 = \{\rho, x_1, x_2, x_3, x_4, u, v\}$ consists of the root ρ, four tree vertices $\{x_1, x_2, x_3, x_4\}$ and two hybrid vertices. However, N' is *not* a galled tree since the subgraph $N'[V^0]$ is a biconnected component that is not a cycle. In other words, N' is a directed acyclic graph that satisfies condition (N1), (N2) and (N3) with respect to the set $\mathcal{L}(N)$, but fails condition (N4).

Let $N = (V, E)$ be a galled tree. As usual in rooted directed acyclic graphs, we say that v is a *child* of u while u is a *parent* of v, for each pair of vertices u and v such that $(u, v) \in E$. A leaf x has, by definition, a unique parent denoted $\text{parent}_N(x)$. If $(u, v) \in E$ and v is a leaf, then v is a *leaf-child* of u. If there is a directed path from u to v (possibly of length zero) in N, then u is an *ancestor* of v while v is a *descendent* of u. We also

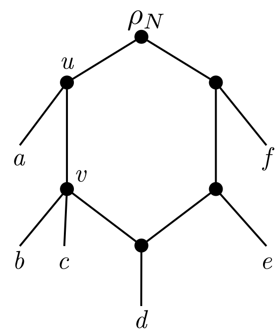


Figure 11: Galled tree that exemplify the notions of ancestors, descendants, least common ancestors and related. See Example 2.11 for details.

denote this relationship by $u \succeq_N v$. If $u \succeq_N v$ but $u \neq v$, we write $u \succ_N v$. In particular, $\succeq_N$ is a partial order on V. We thus say that $x, y \in V$ are *comparable* if $u \succeq_N v$ or $v \succeq_N u$, and *incomparable* otherwise. The latter is occasionally denoted $u \,\|_N\, v$. We let $\mathrm{ancest}_N(v)$ denote the set of ancestors of a vertex v in N, that is,

$$\mathrm{ancest}_N(v) = \{u \in V \mid u \succeq_N v\}.$$

For each subset of vertices $W \subseteq V$ in a galled tree $N = (V, E)$, there is a unique vertex that is an ancestor of all vertices in W, and is minimal with respect to $\succ_N$ among all vertices with that property [18, Lem. 7.9]. This vertex is called the *least common ancestor of* W, and it is denoted by $\mathrm{lca}_N(W)$. For simplicity, we write $\mathrm{lca}_N(x, y)$ for $\mathrm{lca}_N(\{x, y\})$. By definition,

$$\mathrm{lca}_N(x, y) = \min_{\succeq_N} \left(\mathrm{ancest}_N(x) \cap \mathrm{ancest}_N(y)\right).$$

Note that least common ancestors are symmetric, i.e. $\mathrm{lca}(x, y) = \mathrm{lca}(y, x)$. If N is understood from context the subscripts are dropped and we write only $\mathrm{lca}(x, y)$, $\succeq$, $\succ$, $\|$, $\mathrm{parent}(v)$ and $\mathrm{ancest}(v)$.

Example 2.11. Consider the galled tree $N = (V, E)$ on $X = \{a, b, c, d, e, f\}$ in Figure 11. In N, v is a child of u while u is a parent a and v. In particular, a is a leaf-child of u. We also have that u is an ancestor of, among others, a and b. In terms of the order $\succeq_N$ we, for example, have $u \succeq_N v \succeq_N b$. Examples of least common ancestors include $\mathrm{lca}(a, e) = \rho_N$, $\mathrm{lca}(b, c) = v$ and $\mathrm{lca}(a, b) = \mathrm{lca}(c, a) = u$.

Given a vertex v in N, the *subnetwork rooted at* v, denoted $N(v)$, is obtained from the directed graph $N[\{x \mid v \succeq_N x\}]$ by suppressing any of its vertices w such that $\mathrm{indeg}(w) = \mathrm{outdeg}(w) = 1$ and, if its new root v has outdegree one, by deleting the vertex v and its incident edge, so that the new root of $N(v)$ is the child of v. Hence, for any $v \in V$, $N(v)$ will be a galled tree.

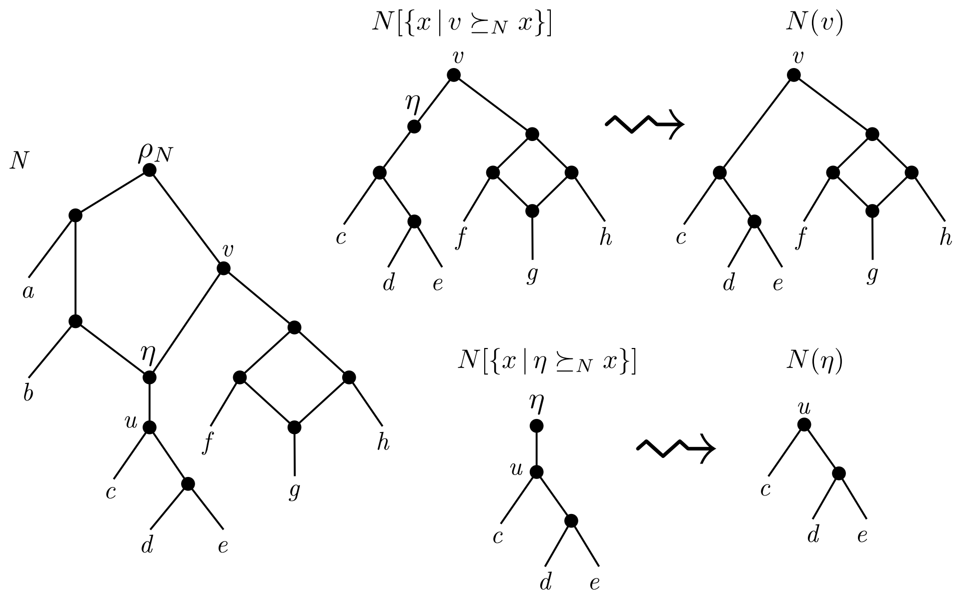


Figure 12: Examples of subnetworks of a given galled tree N (left). The network $N(v)$ is obtained from $N[\{x \,|\, v \succeq_N x\}]$ by suppression of the vertex η (see top row), since η has both in- and outdegree one in $N[\{u \,|\, v \succeq_N u\}]$. The network $N(\eta)$ is obtained from $N\{x \,|\, \eta \succeq_N x\}$ by deletion of η (see bottom row), since the root of $N\{x \,|\, \eta \succeq_N x\}$ has outdegree one.

Example 2.12. Consider the galled tree N in Figure 12, where the two subnetworks $N(u)$ and $N(\eta)$ are depicted as well. To obtain $N(v)$ from $N[\{u \,|\, v \succeq_N u\}]$ one suppresses the vertex η, since both the in- and outdegree of η in $N[\{u \,|\, v \succeq_N u\}]$ is one. Similarly, η is suppressed to obtain $N(\eta)$ from $N[\{u \,|\, \eta \succeq_N u\}]$, since η has outdegree one in the latter network. In other words, $N(\eta) = N(u)$, where $\eta = \text{parent}(u)$.

A galled tree that has no hybrid vertices is a *tree*. An example of a tree is a *caterpillar tree*, whose inner vertices consist of a single (directed) path $P^1 = v_1 v_2 \ldots v_k$ (possibly $k = 1$) such that v_i has a single leaf-child for each $i = 1, 2, \ldots, k-1$ and such that v_k has precisely two children, both which are leafs. Additionally, a caterpillar tree will by definition be rooted at the vertex v_1. Any graph isomorphic to the subgraph of a caterpillar tree induced by v_k and its two children x and y is called a *cherry*. In that case, the leaves x and y are said to be *part of a cherry*.

Example 2.13. A caterpillar tree T is given in Figure 13. In T, the subnetwork $N(v_5)$ is its cherry, while u_5 and u_6 are the leaves that are part of a cherry. Another example of a caterpillar tree is the subnetwork $N(\eta)$ of Figure 12.

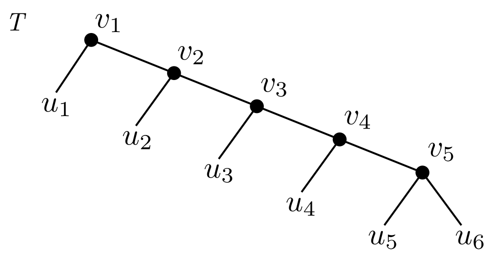


Figure 13: A caterpillar tree T. The leaves u_5 and u_6 are parts of its cherry $T(v_5)$.

A cycle C with sides P^1 and P^2 is *weak* if either (i) P^1 or P^2 has no inner vertex, or if (ii) P^1 and P^2 has precisely one inner vertex each. Weak cycles satisfying (ii) are called *balanced*, those satisfying (i) are called *unbalanced.* The edge from the root to the hybrid in an unbalanced weak cycle is called a *shortcut.* Cycles that are not weak are called *strong.* A galled tree N is *strong* (*weak*) if all its cycles are *strong* (*weak*). Note that trees vacuously are both weak and strong galled trees.

We say that a galled tree N is *elementary* if (i) all its inner vertices are contained in a single cycle C of N, (ii) no child of ρ_N is a leaf and if (iii) each inner vertex $v \neq \rho_N$ has precisely one leaf-child. Note that an elementary galled tree will have a unique hybrid vertex, namely, the unique hybrid vertex of its underlying cycle. Moreover, by condition (i), $\rho_C = \rho_N$. In an elementary galled tree with sides P^1 and P^2, a leaf x whose parent lies on, say, P^1 is said to *belong to* P^1.

Example 2.14. We give three examples of elementary galled trees N, N' and N'' in Figure 14. The underlying cycle of N is unbalanced and weak, while N' has a single balanced, weak cycle. The galled tree N'' is, on the other hand, strong since its underlying cycle is so.

Figure 14 also depicts three galled trees G, G' and G'' that are not elementary. G is not elementary since it is a tree – its inner vertices does not induce a cycle. G' is not elementary since one of its inner vertices has more than one leaf-child, and G'' is not elementary since its root has a leaf-child. Note that G and G' are weak galled trees, and that both G and G'' are strong galled trees.

The galled tree N of Figure 12 constitutes an example of a galled tree which is neither weak nor strong; it has both a (balanced) weak cycle and a strong cycle.

A considerable part of section 3 will study elementary galled trees. We now provide the following simple result for later reference.

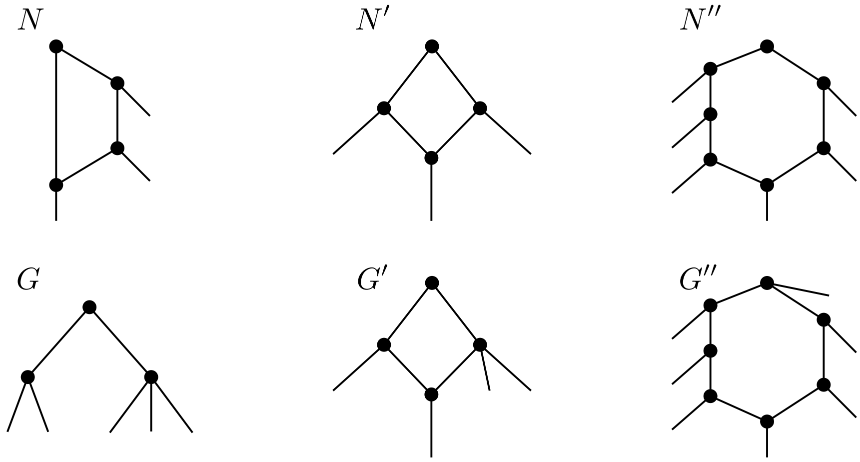


Figure 14: Examples of elementary galled trees (top row) and galled trees that are not elementary (bottom row).

Lemma 2.1. *Let N be an elementary galled tree where v is the child of its hybrid vertex. The following holds for each pair of distinct leaves $x, y \in \mathcal{L}(N)$:*

- *If* $\mathrm{parent}(x)$ *and* $\mathrm{parent}(y)$ *are comparable in N, then* $\mathrm{lca}(x,y) = \max_{\succeq_N}\{\mathrm{parent}(x), \mathrm{parent}(y)\}$. *In particular,* $\mathrm{lca}(v,x) = \mathrm{parent}(x)$ *for each leaf $x \neq v$.*
- *If* $\mathrm{parent}(x)$ *and* $\mathrm{parent}(y)$ *are incomparable in N, then* $\mathrm{lca}(x,y) = \rho_N$.

Proof. Let $N = (V,E)$ be an elementary galled tree whose underlying cycle C has sides P^1 and P^2. Let η denote the unique hybrid vertex of N, so that P^1 and P^2 are disjoint except for the root ρ and η. The unique leaf-child of η is denoted with v. Denote the leaves that belong to P^1 by x_1, x_2, ..., and x_k and those that belong to P^2 by y_1, y_2, ..., and y_l, where k and l are nonnegative integers. Without loss of generality, assume

$$\begin{aligned} \rho \succ_N \mathrm{parent}(x_1) \succ_N \mathrm{parent}(x_2) \succ_N \ldots \succ_N \mathrm{parent}(x_k) = \eta \\ \rho \succ_N \mathrm{parent}(y_1) \succ_N \mathrm{parent}(y_2) \succ_N \ldots \succ_N \mathrm{parent}(y_l) = \eta. \end{aligned} \tag{1}$$

By definition of v we thus have that $v = x_k = y_l$. A direct consequence of the elementary structure of N and (1) is that for any two leaves x and y, their respective parents are comparable (with respect to $\succeq_N$) if and only if x and y belong to the same side of C.

A second implication of (1) and the elementary nature of N is that the set of ancestors of any leaf is easily described; we have

$$\mathrm{ancest}(x_i) = \{\mathrm{parent}(x_i), \mathrm{parent}(x_{i-1}), \ldots, \mathrm{parent}(x_1), \rho\} \tag{2}$$

for each $i \in \{1, \ldots, k-1\}$, and furthermore

$$\text{ancest}(y_i) = \{\text{parent}(y_i), \text{parent}(y_{i-1}), \ldots, \text{parent}(y_1), \rho\} \tag{3}$$

for each $i \in \{1, \ldots, l-1\}$. For v, we have

$$\text{ancest}(v) = V^0(N). \tag{4}$$

To prove the first statement, suppose x and y are leaves such that their respective parents are comparable in N. As discussed earlier, this happens only if x and y belong to the same side of C. By symmetry, we may without loss of generality assume that $x = x_i$ for some $i = 1, \ldots, k-1$ and $y = x_j$ for some $j = i+1, \ldots, k$. By (2) and possibly (4) (the latter if $j = k$), the set of common ancestors of x_i and x_j equals

$$\text{ancest}(x_i) \cap \text{ancest}(x_j) = \text{ancest}(x_i).$$

Taken together with (1), $\text{lca}(x_i, x_j) = \text{parent}(x_i)$ follows. In particular, $\text{lca}(x_i, v) = \text{parent}(x_i)$ holds for each $i \in \{1, \ldots, k-1\}$, since $v = x_k$.

For the second statement, assume x and y are leaves such that $\text{parent}(x)$ and $\text{parent}(y)$ are incomparable in N. Thus x and y belong to different sides of C, in particular $v \notin \{x, y\}$, since v belong to both P^1 and P^2. Hence (2) and (3) implies the set of common ancestors of x and y equals

$$\text{ancest}(x) \cap \text{ancest}(y) = \{\rho\},$$

from which $\text{lca}(x, y) = \rho$ follows immediately.

□

We will consider *labeled* galled trees, that is, a galled tree $N = (V, E)$ on X equipped with a (vertex-)*labeling* $t : V \to \Upsilon$. The *label set* Υ is assumed to contain the element $\odot$, and we assume $t(x) = \odot$ if and only if $x \in X$. This is, however, only a subtle technicality meant to deal with the case when $|V| = 1$, and will not be of any importance in non-trivial cases. When we state that (N, t) is a galled tree, we implicitly assume t to be a labeling of N. Since the least common ancestor of two leaves is uniquely determined, we may define $\mathcal{G}(N, t)$ to be the edge-colored graph with vertex set X such that its coloring map σ satisfies $\sigma(xy) = t(\text{lca}(x, y))$ for each pair of distinct vertices $x, y \in X$. An edge-colored graph Σ is said to be *explained by* (N, t) if $\Sigma \simeq \mathcal{G}(N, t)$. As a special case, Σ is *caterpillar-explainable* if $\Sigma \simeq \mathcal{G}(N, t)$ for a caterpillar tree N.

Example 2.15. We provide a first example of a labeled galled tree (N, t) in Figure 15, alongside the edge-colored graph $(V, \sigma) := \mathcal{G}(N, t)$. For example, the edge between vertices a and e in Σ is colored (dotted) blue, since $t(\text{lca}_N(a, e)) = t(\rho) = \blacktriangle$. Similarly, $t(\text{lca}_N(c, f)) = t(\text{parent}(f)) = \blacksquare$ implies that the edge $\{c, f\}$ is mapped to the (dashed) green edge-color by σ.

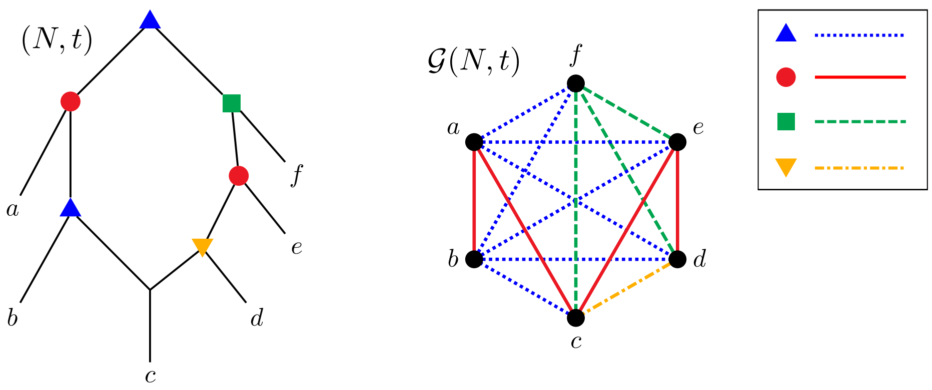


Figure 15: A labeled galled tree (N, t) and the graph $\mathcal{G}(N, t)$ it explains. The legend indicates which labels of N corresponds to what edge colors of $\mathcal{G}(N, t)$.

Remark 2.1. All figures that depict labeled galled trees are drawn "top down", i.e. the edges' direction are from the top-most vertex to the lower vertex. Furthermore, the $\odot$-labels of its leaves are omitted. Most often, three or four distinct labels suffices in examples, and these are indicated with the labels ●, ▲, ■ and ▼. In connection to edge-colors, these labels corresponds to solid red edges, blue dotted edges, dashed green edges and dash-dotted yellow edges, respectively (as shown in the legend in Figure 15). Lastly, the label of hybrid vertices is often not of importance and is thus omitted in some figures; a statement that will be made more precise in Observation 2.1.

A galled tree (N, t) is *discriminating* if, for all adjacent inner vertices u and v, we have $t(u) \neq t(v)$. If $t(u) \neq t(v)$ for all adjacent tree vertices u and v of N, then (N, t) is *quasi-discriminating*. Tree vertices are inner vertices, hence all discriminating galled trees are quasi-discriminating, but not vice versa. However, since trees have no hybrid vertices any quasi-discriminating tree (T, t) is, in fact, discriminating.

Example 2.16. Three galled trees (N, t), (N', t'), and (N'', t'') are given in Figure 16, together with the respective edge-colored graph they explain: $\mathcal{G}(N, t)$, $\mathcal{G}(N', t')$, and $\mathcal{G}(N'', t'')$. One easily verifies that (N, t) is not quasi-discriminating (and thus not discriminating), while (N', t') is quasi-discriminating. The galled tree (N'', t'') is discriminating.

Although tedious, one may verify that $\mathcal{G}(N, t) = \mathcal{G}(N', t')$. On the other hand, $\mathcal{G}(N'', t'') \neq \mathcal{G}(N, t)$, as seen by comparing the color of the edge $\{e, f\}$. Although slightly more complicated to verify formally, it is in fact true that $\mathcal{G}(N'', t'') \not\simeq \mathcal{G}(N, t)$. This hints at the importance of specifying wether a galled tree is (quasi-)discriminating or not; loosely speaking, (N'', t'') is

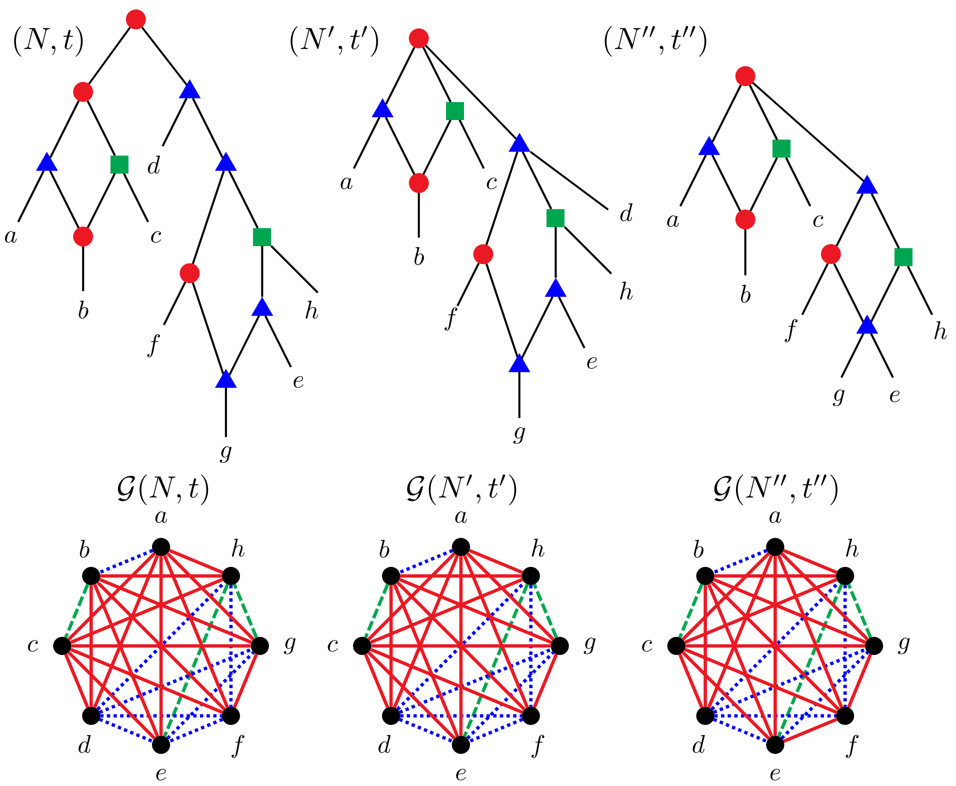


Figure 16: Three galled trees (top row) and the edge-colored graphs they explain (bottom row). See Example 2.16 for details.

obtained from (N, t) by contracting edges whose endpoints are labeled identically, creating a discriminating 'version' of (N, t). But after that process, the two galled trees do not explain the same graph. In contrast, (N', t) can be thought of as being constructed from (N, t) by contracting edges with equal labels on their endpoints, as long as none of the endpoints are hybrid vertices. This only ensures a quasi-discriminating galled tree, but on the other hand, the galled tree that is obtained still explain the same graph.

On a related note, one may observe that there is no vertex $x \in \mathcal{L}(N)$ such that $\mathrm{lca}_N(g, x) = \mathrm{parent}(g)$, nor any $x \in \mathcal{L}(N)$ such that $\mathrm{lca}_N(b, x) = \mathrm{parent}(b)$. Hence the labels of $\mathrm{parent}(g)$ and $\mathrm{parent}(b)$ — the hybrid vertices of N — are, so to speak, superfluous information whilst constructing $\mathcal{G}(N, t)$. This is the reason why we in some places, such as Figure 15, omit the labels of hybrid vertices. Again, we will formalize this further in Observation 2.1.

For a more in-depth discussion about the vaguely introduced concept of edge-contractions in labeled galled trees appearing in Example 2.16, see [19, Sec. 8]. Results in that direction for edge-colored graphs is, unfortunately, out of the scope of this paper but nevertheless worth mentioning.

In Section 4 we will consider when there is a "unique" galled tree that explain a primitive edge-colored graph. We now want to make the concept of uniqueness more precise, since isomorphims as directed graphs is insufficient. Adapting from [22], we say that two (labeled) galled trees (N, t) and (N', t') are *isomorphic* if $\mathcal{L}(N) = \mathcal{L}(N')$ and there exist an isomorphism φ of N and N' (as directed graphs) such that $\varphi(v) = v$ for every $v \in \mathcal{L}(N)$. Furthermore, φ must satisfy

$$t(v) = t(u) \iff t'(\varphi(v)) = t'(\varphi(u))$$

for all $v, u \in V^0_{\text{lca}}(N)$. Here $V^0_{\text{lca}}(N)$ denotes the set of all vertices $v \in V(N)$ for which there exist leaves $x, y \in \mathcal{L}(N)$ such that $v = \text{lca}_N(x, y)$. As leaves are never the least common ancestor of any two leafs, we will have $V^0_{\text{lca}}(N) \subseteq V^0(N)$ for all galled trees N. It is also noteworthy that any isomorphism φ of N and N' will satisfy $\varphi(\rho_N) = \rho_{N'}$, since ρ_N and $\rho_{N'}$ are the only respective vertex with in-degree zero. In symbols we write $(N, t) \simeq (N', t')$ for isomorphic galled trees (N, t) and (N', t'). Furthermore, if (N, t) is stated to be *unique* with respect to satisfying some property $\mathcal{P}$, then we mean that a galled tree (N', t') satisfies $\mathcal{P}$ if and only if $(N, t) \simeq (N', t')$. The definition of isomorphism of labeled galled trees is motivated by the following lemma.

Lemma 2.2. *If (N, t) and (N', t') are labeled galled trees such that $(N, t) \simeq (N', t')$, then the identity map is an isomorphism of $\mathcal{G}(N, t)$ and $\mathcal{G}(N', t')$.*

Proof. Let φ be an isomorphism of the galled trees (N, t) and (N', t'). Since $V := V(\mathcal{G}(N, t)) = \mathcal{L}(N) = \mathcal{L}(N') = V(\mathcal{G}(N', t'))$, the identity map on $\binom{V}{2}$ is indeed a bijective function that maps each edge of $\mathcal{G}(N, t)$ to an edge of $\mathcal{G}(N', t')$. Denote the edge-coloring of $\mathcal{G}(N, t)$ and $\mathcal{G}(N', t')$ by σ respectively σ', so that $\mathcal{G}(N, t) = (V, \sigma)$ and $\mathcal{G}(N', t') = (V, \sigma')$. Note that $\sigma(\{x, y\}) = \sigma(\{u, v\})$ for edges $\{x, y\}, \{u, v\} \in \binom{V}{2}$ if and only if $t(\text{lca}_N(x, y)) = t(\text{lca}_N(u, v))$, by definition of $\mathcal{G}(N, t)$. Furthermore, by definition of the isomorphism φ we have $t(\text{lca}_N(x, y)) = t(\text{lca}_N(u, v))$ if and only if $t'(\varphi(\text{lca}_N(x, y))) = t'(\varphi(\text{lca}_N(u, v)))$. Since φ is an isomorphism of N and N' as directed graphs and φ fixates all leaves we have $t'(\varphi(\text{lca}_N(x, y))) = t'(\varphi(\text{lca}_N(u, v)))$ if and only if

$$\begin{aligned} t'(\text{lca}_{N'}(x, y)) &= t'(\text{lca}_{N'}(\varphi(x), \varphi(y))) \\ &= t'(\text{lca}_{N'}(\varphi(u), \varphi(v))) = t'(\text{lca}_{N'}(u, v)). \end{aligned}$$

Lastly, by definition of $\mathcal{G}(N', t')$ we have $t'(\text{lca}_{N'}(x, y)) = t'(\text{lca}_{N'}(u, v))$ if and only if $\sigma'(\{x, y\}) = \sigma'(\{u, v\})$. We have thus shown that

$$\sigma(\{x, y\}) = \sigma(\{u, v\}) \iff \sigma'(\{x, y\}) = \sigma'(\{u, v\})$$

for all $\{x, y\}, \{u, v\} \in \binom{V}{2}$. That is, $\mathcal{G}(N, t)$ and $\mathcal{G}(N', t')$ are isomorphic via the identity map. $\square$

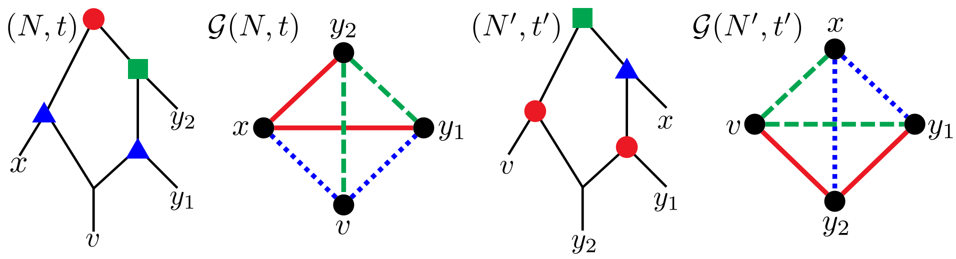


Figure 17: Example where $\mathcal{G}(N,t) \simeq \mathcal{G}(N',t')$, although $(N,t) \not\simeq (N',t')$.

Although $(N,t) \simeq (N',t')$ implies $\mathcal{G}(N,t) \simeq \mathcal{G}(N',t')$, the reverse implication need not hold. An example is in place.

Example 2.17. Consider the two galled trees (N,t), (N',t') and the respective edge-colored graph they explain in Figure 17. It is not complicated to see that $\mathcal{G}(N,t)$ and $\mathcal{G}(N',t')$ are isomorphic; in fact, the identity map on $\binom{V}{2}$ for $V = \{x, v, y_1, y_2\}$ is an explicit isomorphism. However, although there is an isomorphism $\varphi : N \to N'$ of N and N' as directed graphs, such a map must necessarily satisfy $\varphi(v) = y_2$. Hence no isomorphism of N and N' may be extended to an isomorphism of (N,t) and (N',t').

We have already mentioned that the label of certain hybrid vertices is, at times, redundant. A more precise statement is the following.

Observation 2.1. As a direct consequence of Lemma 2.1 we have that if (N,t) is an elementary galled tree with hybrid vertex η, then $V^0_{\mathrm{lca}}(N) = V^0(N) \setminus \{\eta\}$.

Now that we have introduced labeled galled trees, we can also motivate why weak cycles are called weak, in the following lemma. A stronger, but highly related statement is given in [19, Lemma 5.4].

Lemma 2.3. *Let (N,t) be an elementary galled tree with underlying cycle C. If one of the following two conditions is satisfied*

- *C is an unbalanced weak cycle, or*
- *C is a balanced weak cycle and $\mathcal{G}(N,t)$ has at most two edge-colors*

then there is a tree (T',t') such that $\mathcal{G}(T',t') = \mathcal{G}(N,t)$.

Proof. First assume C is an unbalanced weak cycle of the elementary galled tree (N,t). That is, C contains the shortcut (ρ,η), where ρ is the root of C (and indeed of N) and η is its unique hybrid vertex. By definition of elementary, ρ has two children. One of them is η, and we denote the other with v. Neither child is a leaf, since N is elementary. Consider the subnetwork $(N',t') := N(v)$, which clearly is a tree. Since ρ is the only vertex of N that either precedes v or is incomparable to v with respect to $\succeq_N$, we have $\mathcal{L}(N(v)) = \mathcal{L}(N)$. Moreover, there are only two vertices that appear in N but not in N': ρ and η, the former since $\rho \succ_N v$ and the latter since η is suppressed in $N(v)$. By Lemma 2.1 there are no two leaves x and y of N such that $\mathrm{lca}_N(x, y) \in \{\rho, \eta\}$. Thus $\mathrm{lca}_N(x, y) = \mathrm{lca}_{N'}(x, y)$ for all leaves x and y, from which $\mathcal{G}(N', t') = \mathcal{G}(N, t)$ follows.

Secondly, assume C is a balanced weak cycle such that $\mathcal{G}(N, t)$ has at most two edge-colors. Then N has four inner vertices denoted ρ, x', y' and η. The cycle C has four edges: (ρ, x'), (ρ, y'), (x', η) and (y', η). Moreover, N has three leafs: x as a child of x', z as a child of η and y as a child of y'. We now enumerate all possible labelings t of N where $\mathcal{G}(N, t)$ has at most two edge-colors. Note that since there are no two leaves u and v such that $\mathrm{lca}(u, v) = \eta$, the label of η is irrelevant. We may, for argument's sake, fix the label of ρ, say $t(\rho) = \bullet$. We only have three possibilities for the labels $t(x')$ and $t(y')$, corresponding to four galled trees:

(i) Both x' and y' have the same label as ρ i.e. $t(x') = t(y') = \bullet$.

(ii) Neither x' nor y' have the same label as ρ i.e. $t(x') = t(y') = \blacktriangle \neq \bullet$.

Precisely one of x' and y' has the same label as ρ i.e.

(iii) $t(x') = \blacktriangle \neq \bullet = t(y')$ or

(iv) $t(x') = \bullet \neq \blacktriangle = t(y')$.

We depict the four possible labeled galled trees $(N, t_{(i)})$, $(N, t_{(ii)})$, $(N, t_{(iii)})$, and $(N, t_{(iv)})$ in Figure 18, alongside the respective edge-colored graph they explain. Figure 18 also depicts a tree $(T_{(\ell)}, t'_{(\ell)})$ for each $\ell \in \{i, ii, iii, iv\}$ such that $\mathcal{G}(T_{(\ell)}, t'_{(\ell)}) = \mathcal{G}(N, t_{(\ell)})$. In other words, if C is a balanced weak cycle and $\mathcal{G}(N, t)$ has two edge-colors, then there exist a tree (T', t') such that $\mathcal{G}(T', t') = \mathcal{G}(N, t)$. $\square$

2.4 Modules and modular decomposition

The central definition of this section is that of a *module* of an edge-colored graph $\Sigma = (V, \sigma)$: it is a subset $M \subseteq V$ such that for all $x, y \in M$ and all $z \notin M$ we have $\sigma(xz) = \sigma(yz)$. In other words, for each $z \notin M$ there is a single color k such that for every $x \in M$ we have $\sigma(xz) = k$. Note, however, that we might still have $\sigma(zx) \neq \sigma(z'x)$ for some $x \in M$ and distinct $z, z' \notin M$.

Example 2.18. Consider the edge-colored graph $\Sigma = (V, \sigma)$ in Figure 19. Clearly,

$$\Sigma = \Sigma[\{a, b, c\}] \otimes_{\blacksquare} \Sigma[\{d, e, f, g\}],$$

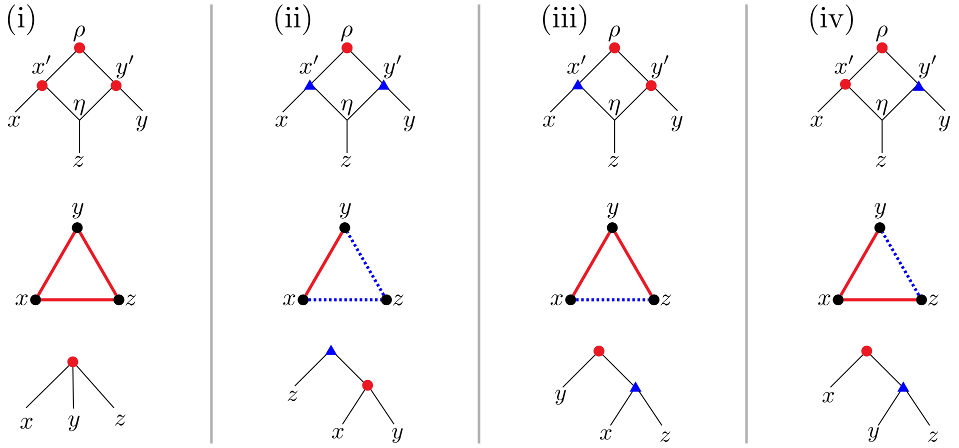


Figure 18: The top row depicts all possible elementary galled trees (N, t) such that its underlying weak cycle is balanced and $\mathcal{G}(N, t)$ has at most two edge-colors. The middle row depicts the respective edge-colored graph the galled tree above explains. The bottom row depicts trees that explain the edge-colored graph above it. In other words, for each galled tree (N, t) in the top row, there exists a tree (T, t') such that $\mathcal{G}(N, t) = \mathcal{G}(T, t')$, see Lemma 2.3.

so we have $\sigma(xy) = \sigma(x'y')$ for every set of vertices such that $x, x' \in \{a, b, c\}$ and $y, y' \in \{d, e, f, g\}$. Hence both $\{a, b, c\}$ and $\{d, e, f, g\}$ are modules of Σ. The sets $\{a, b\}$, $\{d, e\}$ and $\{d, e, f\}$ may also be verified to be modules. The set $\{e, f\}$ is not a module, since $\sigma(d, e) = \blacktriangle \neq \bullet = \sigma(d, f)$.

Modules in 2-structures where introduced under the name *clans* in [8]. We use the word module to emphasize the underlying connection to the more graph theoretic notion of a module, as in e.g. [22]. Modules are also known as clumps [2], autonomous sets [27], and stable sets [28], to only mention a few. Note that in any edge-colored graph $\Sigma = (V, \sigma)$, the empty set, all of V and the singleton sets $\{v\}$ for $v \in V$ always are modules. These modules are called *trivial*, and all other modules *nontrivial*. To avoid technicalities, we assume all modules in this contribution are nonempty, unless otherwise stated. We let $\mathbb{M}(\Sigma)$ denote the set of all modules of Σ. If $\mathbb{M}(\Sigma)$ consists of trivial modules only, then Σ is a *primitive* (edge-colored) graph. In particular all edge-colored graphs on one or two vertices are primitive.

Two modules M and M' are said to *overlap* if they have a nonempty intersection and one is not contained in the other, that is, if $M \cap M' \notin \{\emptyset, M, M'\}$. A module M is *strong* if M does not overlap with any other module. We let $\mathbb{M}_{\text{str}}(\Sigma) \subseteq \mathbb{M}(\Sigma)$ denote the set of strong modules of an edge-colored graph Σ. Since the trivial modules of Σ are strong, we have $\mathbb{M}_{\text{str}}(\Sigma) \neq \emptyset$.

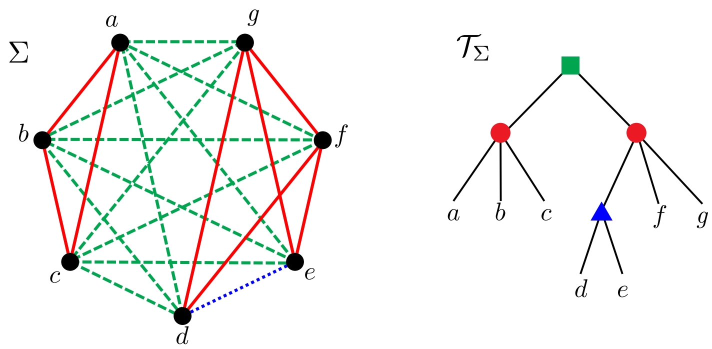


Figure 19: An edge-colored graph Σ (left) with its modular decomposition tree $(\mathcal{T}_\Sigma, \tau_\Sigma)$ (right).

Example 2.19. For the edge-colored graph $\Sigma = (V, \sigma)$ of Figure 19, one may verify that $\mathbb{M}_{\text{str}}(\Sigma)$ contains precisely the trivial modules of Σ, together with the modules $\{a, b, c\}$, $\{d, e, f, g\}$ and $\{d, e\}$. Σ has many other modules, for example the module $\{d, e, f\}$ is not strong, since it overlaps with the module $\{f, g\}$; we have $\{d, e, f\} \cap \{f, g\} = \{f\}$.

A strong module M of an edge-colored graph Σ is called *k-series* if there exist a partition $\{M', M''\}$ of M such that $\Sigma[M] = \Sigma[M'] \otimes_k \Sigma[M'']$. Although the partition $\{M', M''\}$ is not necessarily unique, the color k is.

Lemma 2.4. *If M is a k-series module and a k'-series module, then $k = k'$.*

Proof. Let $\Sigma = (V, \sigma)$ be an edge-colored graph and suppose M is a module of Σ that is simultaneously k-series and k'-series. Let $\{X, X'\}$ and $\{Y, Y'\}$ be partitions of M such that

$$\begin{aligned} \Sigma[M] &= \Sigma[X] \otimes_k \Sigma[X'] && (5) \\ &= \Sigma[Y] \otimes_{k'} \Sigma[Y']. && (6) \end{aligned}$$

First consider the case where $X \cap Y = \emptyset$. Then we must have $X \subseteq Y'$ and $Y \subseteq X'$, since $M = X \cup X' = Y \cup Y'$. For any $x \in X$ and any $y \in Y$ we thus have $\sigma(xy) = k$ by (5) and $\sigma(xy) = k'$ by (6). Symmetric arguments can be made for the case when $X' \cap Y' = \emptyset$. We may thus assume the existence of $x \in X \cap Y$ and $y \in X' \cap Y'$. But then, again, (5) ensures that $\sigma(xy) = k$ and (6) that $\sigma(xy) = k'$. Therefore, $k = k'$. □

Clearly, there may not be any color k such that a given strong module is k-series (consider, for example, singleton modules). If so, then the module in question is called *prime*.

In what follows, let $\Sigma = (V, \sigma)$ be an arbitrary edge-colored graph with edge-colors Υ. The set $\mathbb{M}_{\text{str}}(\Sigma)$ of strong modules may be represented uniquely by a tree $\mathcal{T}_\Sigma$, known as the *modular decomposition tree* (MDT) of Σ, constructed as follows [22]. $\mathcal{T}_\Sigma$ is the tree with vertex set $\mathbb{M}_{\text{str}}(\Sigma)$ and edges (M, M') for each $M, M' \in \mathbb{M}_{\text{str}}(\Sigma)$ such that $M' \subsetneq M$ and such that there is no $X \in \mathbb{M}_{\text{str}}(\Sigma)$ with $M' \subsetneq X \subsetneq M$. The root of $\mathcal{T}_\Sigma$ is, by definition, the vertex V. By construction, we will in particular have that $\mathcal{L}(\mathcal{T}_\Sigma)$ consists of the singleton subsets of V, and we can thus identify the leaves of $\mathcal{T}_\Sigma$ with the vertices of Σ.

The tree $\mathcal{T}_\Sigma$ encodes the structure of the strong modules of Σ with respect to (non-)inclusion. Even more can be said when equipping $\mathcal{T}_\Sigma$ with the following *labeling* $\tau_\Sigma : \mathbb{M}_{\text{str}}(\Sigma) \to \Upsilon \cup \{\textsc{Prime}, \odot\}$, defined by

$$\tau_\Sigma(M) = \begin{cases} k & \text{if } M \text{ is } k\text{-series} \\ \textsc{Prime} & \text{if } |M| > 1 \text{ and } M \text{ is prime} \\ \odot & \text{otherwise i.e. if } |M| = 1. \end{cases}$$

Note that τ_Σ is well-defined by Lemma 2.4. The pair $(\mathcal{T}_\Sigma, \tau_\Sigma)$ at least partially explains the underlying graph Σ, in the sense that if $\tau_\Sigma(\text{lca}_{\mathcal{T}_\Sigma}(x, y)) = k \neq \textsc{Prime}$ for two distinct $x, y \in V$ then we have $\sigma(x, y) = k$. If, on the other hand, $\tau_\Sigma(\text{lca}_{\mathcal{T}_\Sigma}(x, y)) = \textsc{Prime}$, then $\mathcal{T}_\Sigma$ reveal no information about the color of the edge $\{x, y\}$. In other words, the pair $(\mathcal{T}_\Sigma, \tau_\Sigma)$ is a labeled tree in the sense of Section 2.3 if and only if there is no internal node v of $\mathcal{T}_\Sigma$ such that $\tau_\Sigma(v) = \textsc{Prime}$. In that case, $\mathcal{G}(\mathcal{T}_\Sigma, \tau_\Sigma) = \Sigma$. If, on the other hand, there is some v such that $\tau_\Sigma(v) = \textsc{Prime}$, then τ_Σ is not a valid labeling of $\mathcal{T}_\Sigma$, so that $\mathcal{G}(\mathcal{T}_\Sigma, \tau_\Sigma)$ is undefined.

The modular decomposition tree and its canonical labeling is rarely straightforward to compute by hand, but it can be computed in $O(n^2)$ time for an edge-colored graph with n vertices [10]. For graphs without edge-colors (that is, edge-colored with two colors), computation of the MDT is even faster, namely with a time complexity of $O(n + m)$ for graphs with n vertices and m edges. Several such algorithms exist: see e.g. [5, 26, 29]. For an overview on the topic of computing modular decomposition, see [15].

Example 2.20. Consider Figure 20, in which an edge-colored graph $\Sigma = (V, \sigma)$ is given. Its MDT $\mathcal{T} := \mathcal{T}_\Sigma$ with labeling $\tau := \tau_\Sigma$ is depicted much like the galled trees, except Prime-labels are indicated with a circled P. Note that for each strong module M in

$$\mathbb{M}_{\text{str}}(\Sigma) = \{\{a\}, \{b\}, \ldots, \{h\}, \{a, b, c\}, \{e, f, g, h\}, \{d, e, f, g, h\}, V\},$$

there exist a corresponding vertex v of $\mathcal{T}$ such that M consists of the leaves which are a descendent of (or equal to) the vertex v. As seen by inspecting the subgraphs $\Sigma[M]$ for each $M \in \mathbb{M}_{\text{str}}$, Σ contains both non-trivial prime modules ($\{a, b, c\}$ and $\{e, f, g, h\}$) and k-series modules. More precisely, $\{d, e, f, g, h\}$ is a ▲-series module while V is a ●-series module. Another point worth noting is that $(\mathcal{T}, \tau)$ contains no ■-labels even though e.g.

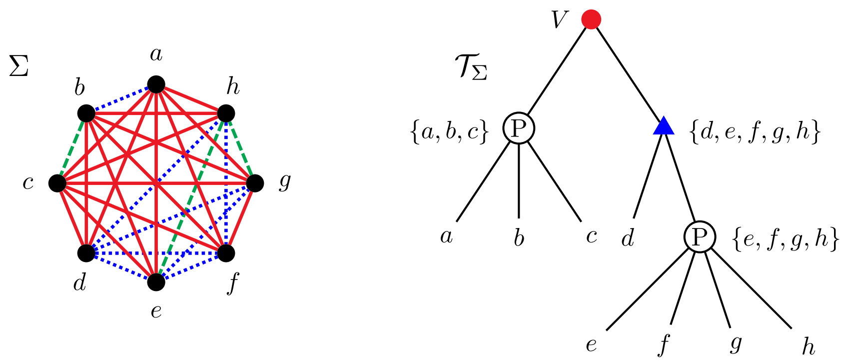


Figure 20: An edge-colored graph Σ (left) alongside its MDT (right). In contrast to the edge-colored graph in Figure 19, Σ contains both non-trivial prime modules as well as two k-series modules.

$\sigma(bc) =$ ■. It is thus evident that Σ cannot be reconstructed solely from the information in $(\mathcal{T}, \tau)$.

For an example of an edge-colored graphs whose MDT contains no Prime-labeled vertices, see Figure 19.

At this point it is worth recalling that this thesis aims at providing crucial results needed to characterize which edge-colored graphs can be explained by galled trees. Since trees are galled trees, partial results are already available. To be more precise, it is obvious that if $(\mathcal{T}_\Sigma, \tau_\Sigma)$ contains no Prime-labeled inner vertices (a condition equivalent to Σ having no non-trivial prime modules), then there is a tree that explains Σ — namely $(\mathcal{T}_\Sigma, \tau_\Sigma)$ itself. In the context of MDTs with no Prime-labeled vertices, a class of particularly interesting graphs are the *uniformly non-primitive* ($\mathfrak{unp}$) edge-colored graphs, that is, edge-colored graphs that has no (induced) primitive subgraph with three or more vertices.

Another important line of research in connection to this thesis is the study of cographs, a class of graphs (with no edge-colors) that can be characterized in many ways (see e.g. [22, Thm. 2.3]); for us it suffices to know that a *cograph* is a graph (without edge-colors) that contain no induced P_4 [4]. They are of some importance for us, since they are the only graphs without edge-colors that can be explained by a tree[1].

Before summarizing existing results on edge-colored graphs that can be explained by trees, we also need to introduce a small yet significant edge-colored graph: the *rainbow triangle*. It is an edge-colored graph on three vertices where its three edges are distinctly colored. With that, we formulate the following theorem.

[1]Technically speaking, we have not introduced the concept of trees that explain graphs without edge-colors, although it is done completely analogous to Section 2.3. To be formal, we can instead state that the cographs are the only graphs whose reinterpretation as an edge-colored graph with two colors can be explained by a tree.

Theorem 2.5. *Let Σ be an edge-colored graph. The following statements are equivalent.*

(1) Σ is explained by a tree.

(2) $\mathbb{M}_{str}(\Sigma)$ contains no non-trivial prime modules.

(3) $(\mathcal{T}_\Sigma, \tau_\Sigma)$ contains no Prime*-labeled inner vertices and thus explains Σ.*

(4) Σ is $\mathfrak{unp}$.

(5) Σ has no primitive (induced) subgraph with three or four vertices.

(6) Σ satisfies both

(a) For each edge-color k, the monochromatic subgraph $\Sigma|_k$ is a cograph, and

(b) Σ contains no rainbow triangle.

Proof. For (1) $\implies$ (2), let (T, t) be a tree and consider the edge-colored graph $(V, \sigma) := \mathcal{G}(T, t)$. Suppose it contains a nontrivial strong module M, that is, a module such that $1 < |M| < |V|$. Let v be the least common ancestor of the vertices of M. If v has a leaf-child x that is contained in M, then put $M' = \{x\}$. Otherwise, let $M' = \mathcal{L}(N(u)) \cap M$ for some child u of v such that $\mathcal{L}(N(u)) \cap M \neq \emptyset$. Lastly, put $M'' = M \setminus M'$. Note that $M', M'' \neq \emptyset$ and that, by construction

$$\sigma(xy) = t(\mathrm{lca}_T(x, y)) = t(v) = t(\mathrm{lca}_T(x', y')) = \sigma(x'y')$$

for all $x, x' \in M'$ and all $y, y' \in M''$. In other words, M is a k-series module, where $k = t(v)$. Hence M is not a prime module, and since M was chosen arbitrarily, condition (2) holds. Clearly, (2) $\implies$ (3) $\implies$ (1) follows directly from the respective definitions. The equivalences of statements (3) through (5) is provided in [11, Thm. 3.6], where the equivalences are proved for general 2-structures (for which edge-colored graphs are a special case). Similarly, the equivalence (5) $\iff$ (6) is provided for 2-structures in [22, Thm. 3]. ☐

With Theorem 2.5 in mind, we can thus conclude that the edge-colored graph in Figure 19 is $\mathfrak{unp}$, since it is explained by its MDT. On the other hand, the edge-colored graph Σ in Figure 20 is not $\mathfrak{unp}$, since its MDT contains Prime-labeled inner vertices.

For later reference, we now consider edge-colored graphs on three vertices. There are, up to isomorphism, precisely three possibilities: rainbow triangles, graphs with the three edges in the same color and, lastly, graphs with two edges in one color and the third edge in a second color. We wish to investigate these three (isomorphism classes) of edge-colored graphs closer.

First, let $\Sigma = (V, \sigma)$ be a rainbow triangle. Since each of its vertices is incident to two distinctly colored edges it has no non-trivial modules and is thus a primitive graph. By Theorem 2.5 this means that there is no tree that explains Σ. Moreover, since V is a prime module, the MDT of Σ consists of a root labeled Prime with three leaf-children. However, as seen in Figure 21, there exist a (elementary) galled trees that explains Σ.

Secondly, if $\Sigma' = (V, \sigma')$ is an edge-colored graph on three vertices that satisfies $\sigma'(xy) = k$ for each distinct $x, y \in V$ and some fixed color k, then any subset of V is a module, which implies that only the trivial modules are strong. Hence $\mathcal{T}_{\Sigma'}$ consists of a root with three leaf-children, just like the MDT of a rainbow triangle. However, V is a k-series module, so $\tau_{\Sigma'}$ labels the root of $\mathcal{T}_{\Sigma'}$ with the color k. That is, $(\mathcal{T}_{\Sigma'}, \tau_{\Sigma'})$ explains Σ'.

Lastly, let $\Sigma'' = (V, \sigma'')$ be an edge-colored graph on three vertices a, b and c that satisfies $\sigma''(ab) = \sigma''(ac) = k$ and $\sigma(bc) = l \neq k$ for two colors k and l. One may verify that $(\mathcal{T}_{\Sigma''}, \tau_{\Sigma''})$ (1) has a root labeled k, (2) the root has two children: the leaf a and an internal vertex labeled l, and (3) that internal vertex has two leaf-children, namely b and c. The tree $(\mathcal{T}_{\Sigma''}, \tau_{\Sigma''})$ explains Σ''. See Figure 21 for the respective MDTs of Σ' and Σ''.

Since edge-colored graphs with only one or two vertices easily are verified to be explained by their respective MDT, one conclusion here is that a rainbow triangle is the smallest edge-colored graph that is not explained by a tree. It is also the smallest non-trivial primitive edge-colored graph. We collect the discussion about small edge-colored graphs in the following observation.

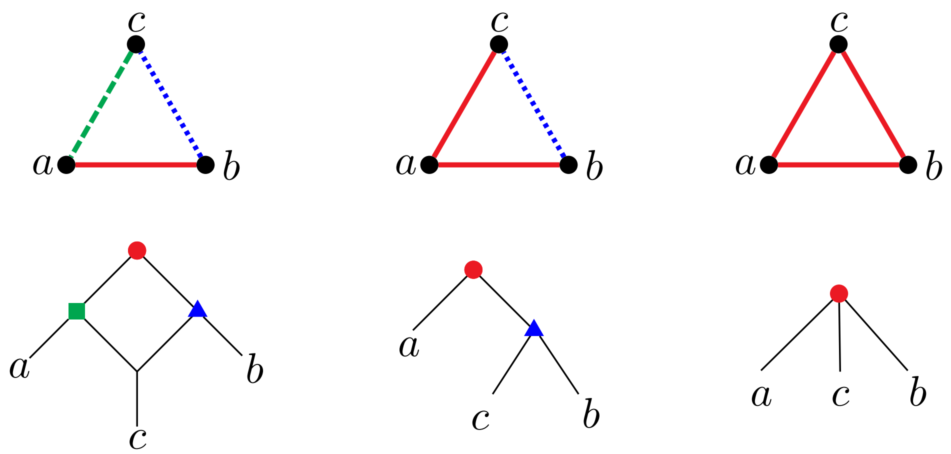


Figure 21: The three possible edge-colored graphs on three vertices are given in the top row. The bottom row depicts three (galled) trees that explain the corresponding edge-colored graph above it. Note in particular that the middle and rightmost edge-colored graphs are explained by their respective MDT, which coincides with the given trees.

Observation 2.2. The smallest edge-colored graph that is not explained by a tree but by a galled tree is the rainbow triangle. In other words, every other edge-colored graph with at most three vertices is explained by a tree. Moreover, rainbow triangles are primitive edge-colored graphs, and there are no other edge-colored graphs with three vertices that are primitive.

To reconnect to the main topic of this section there are, so to speak, two 'extremes' when it comes to how much information of the underlying edge-colored graph that can be reestablished from its MDT. On one hand, $\mathfrak{unp}$ edge-colored graphs are explained by their MDT. On the other hand, we have primitive graphs. The trivial primitive graphs on one and two vertices are admittedly explained by their MDT, but for primitive graphs with at least three edge colors the MDT will, by definition, consist of a root with leaf-children only. Moreover, its root will always be PRIME-labeled since if not, then every edge of the graph would have the same color, which contradicts the primitivity. Hence for these graphs, one can not reestablish any edge's color by inspecting the MDT. Not all edge-colored graphs are covered by these two cases, of course; see for example the edge-colored graph Figure 20, which is neither primitive nor $\mathfrak{unp}$. Since each edge-colored graph that is not $\mathfrak{unp}$ contains some primitive subgraph with at least three vertices, we are motivated to study which primitive edge-colored graphs can be explained by a galled tree.

3 Galled trees and primitive edge-colored graphs

Recall that an edge-colored graph is primitive if it has only one or two vertices, or if its MDT has a Prime-labeled root with leaf-children only. Since primitive graphs in general are not explained by their MDT — in fact by no tree — it would be helpful to instead explain them with the more generalized notion of galled trees. Hence the goal of this section is to establish a characterization of primitive edge-colored graphs that can be explained by galled trees. In the first subsection, we introduce so-called polar-cats, state the main theorem and aim at giving an intuitive understanding of its proof. We then provide a plethora of structural results and combine these to a proof of the characterization.

3.1 Definition, examples and main statement

We begin by considering labeled, elementary and quasi-discriminating galled trees. What type of edge-colored graph do they explain? How can we characterize these graphs? Although not trees, elementary galled trees have a lot of 'local' structure that resembles caterpillar trees. As we will see, this means that the edge-colored graphs they explain can, so to say, be split into two parts that have a very particular structure. Additionally, these edge-colored graphs are, vaguely speaking, primitive yet *almost* $\mathfrak{unp}$; a single vertex makes the difference. To be precise, we introduce the following definition.

Definition 3.1. We say that an edge-colored graph $\Sigma = (V, \sigma)$ is a *polar-cat* if there exists a vertex $v \in V$ and two induced subgraphs $\Omega_1 = (W_1, \omega_1)$ and $\Omega_2 = (W_2, \omega_2)$ of Σ such that

(i) $|W_1| \geq 2$, $|W_2| \geq 2$, $W_1 \cup W_2 = V$ and $W_1 \cap W_2 = \{v\}$

(ii) $\Sigma - v = (\Omega_1 - v) \otimes_k (\Omega_2 - v)$ for some color k.

(iii) For $i \in \{1, 2\}$, Ω_i can be explained by a labeled discriminating caterpillar tree (T_i, t_i) such that

  (a) v is part of its cherry, and such that

  (b) $t_i(\rho_{T_i}) \neq k$.

In that case, we also say that Σ is a (v, Ω_1, Ω_2)-polar-cat. The color k is called its *join-color*.

The concept of polar-cats was introduced in [19] in the context of graphs (without edge-colors) explained by galled trees, and the definition above is a direct generalization to edge-colored graphs. The name polar-cat stems from the subgraphs being **polar**izing, i.e. the requirement of $t_i(\rho_{T_i}) \neq k$, and **cat**erpillar-explainable.

Example 3.1. Consider the edge-colored graph $\Sigma = (V, \sigma)$ in Figure 22. We will verify that Σ is a (c, Ω_1, Ω_2)-polar-cat for $\Omega_1 := \Sigma[\{a, b, c\}]$ and $\Omega_2 := \Sigma[\{c, d, e, f\}]$. By definition, it holds that $|V(\Omega_1)| = 3 \geq 2$, $|V(\Omega_2)| = 4 \geq 2$, $V = V(\Omega_1) \cup V(\Omega_2)$ and $V(\Omega_1) \cap V(\Omega_2) = \{c\}$, so condition (i) holds.

For condition (ii), it is easy to verify that

$$\Sigma - c = (\Omega_1 - c) \otimes_{\blacktriangle} (\Omega_2 - c),$$

see Figure 22. That is, the join-color of Σ is $\blacktriangle$. The figure also provides two caterpillar trees (T_1, t_1) and (T_2, t_2) — it is left to the reader to verify that they explain Ω_1 and Ω_2, respectively. Note that c is part of the cherry of both T_1 and T_2, so (iii)(a) is satisfied. Moreover,

$$t_1(\rho_{T_1}) = \bullet \neq \blacktriangle \quad \text{and} \quad t_2(\rho_{T_2}) = \blacksquare \neq \blacktriangle,$$

which verifies (iii)(b). Altogether, conditions (i)-(iii) are satisfies, so Σ is indeed a polar-cat.

For later reference we note the following.

Observation 3.1. By condition (i) of Definition 3.1 a polar-cat must have at least three vertices. Considering the same three possible edge-colored graphs on three vertices as in Observation 2.2, one notes that a rainbow triangle $\Sigma = (V, \sigma)$ with $V = \{a, b, c\}$ is, for example, an $(a, \Sigma[\{a, b\}], \Sigma[\{a, c\}])$-polar-cat. Moreover, neither of the other two (isomorphism classes of) edge-colored graphs are polar-cats, since any choice of the vertex v and the subgraphs Ω_1 and Ω_2 will fail condition (iii)(b).

We now state the main result of this section, although we leave the proof for later.

Theorem 3.2. *Let Σ be an edge-colored graph. The following statements are equivalent.*

(a) Σ is either a rainbow triangle or is explained by a strong elementary and quasi-discriminating galled tree.

(b) Σ is a polar-cat.

(c) Σ is primitive, has at least three vertices and is explained by a galled tree.

For the reader's convenience, we recall that an elementary galled tree has precisely one underlying cycle, where each internal node excluding the root has precisely one leaf-child. It is strong whenever that cycle is strong i.e. both sides have at least one inner vertex and at most one side has only one inner vertices. A galled tree is quasi-discriminating if adjacent internal vertices, except possibly hybrid vertices, are labeled distinctly.

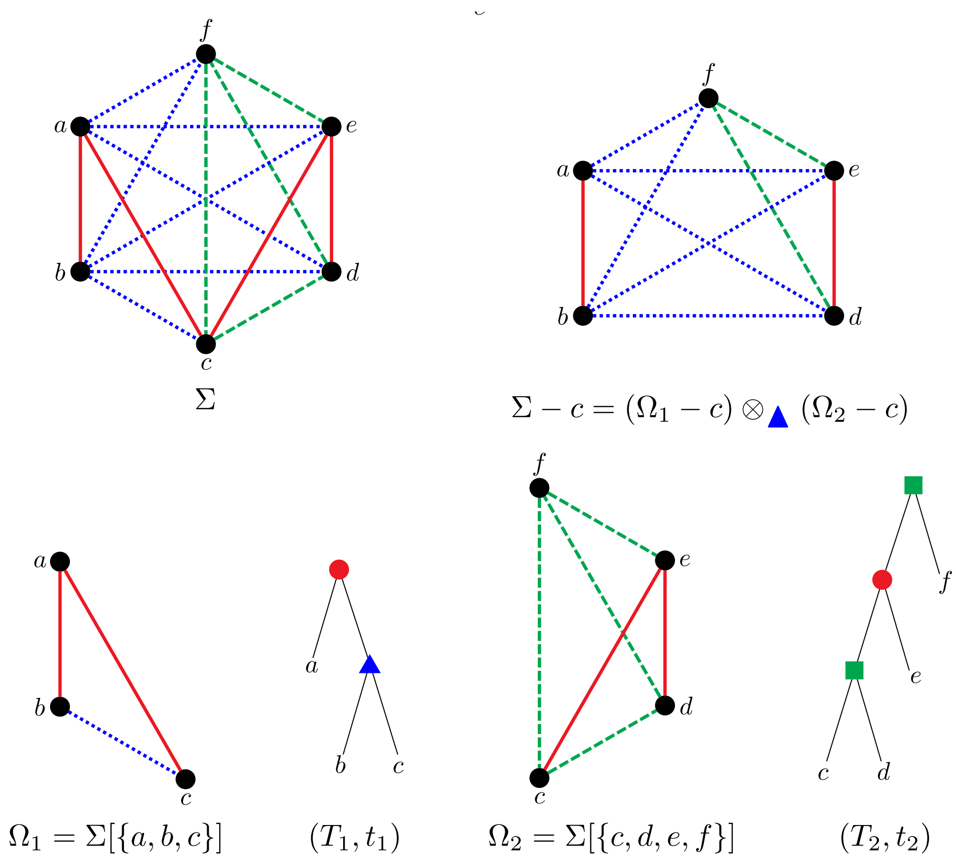


Figure 22: We provide an edge-colored graph Σ which is a (c, Ω_1, Ω_2)-polar-cat, where $\Omega_1 = \Sigma[\{a, b, c\}]$ and $\Omega_2 = \Sigma[\{c, d, e, f\}]$. The subgraphs Ω_1, Ω_2 and $\Sigma - c$ are also depicted. Moreover the trees (T_1, t_1) and (T_2, t_2) are given alongside the respective subgraph they explain.

It is worth emphasizing that Theorem 3.2 is best thought of as a characterization of primitive graphs that can be explained by a galled tree; the structure of polar-cats is most important in connection to the theorem's first statement. In fact, it is not too hard to get an intuitive understanding of why the first two statements are equivalent: we outline the ideas in the following example.

Example 3.2. Again consider the polar-cat Σ in Figure 22 and Example 3.1. In Figure 23 we show how the trees (T_1, t_1) and (T_2, t_2) can be turned into an elementary galled tree. First the two trees are joined under a common root, which is labeled with the join-color of Σ. Secondly the two different occurrences of the leaf c, which appear in the cherry of the respective tree, are merged and brought under a common hybrid vertex. The resulting elementary galled tree (N, t) will explain Σ, since (N, t), loosely speaking, inherits a lot of 'local' structure from (T_1, t_1) and (T_2, t_2). More precisely, we for example have that $\text{lca}_T(x, y) = \text{lca}_{T_i}(x, y)$ for each pair of distinct $x, y \in V(\Omega_i) \setminus \{c\}$ (for both $i = 1$ and $i = 2$). The other technical aspects of the definition of polar-cats can also be motivated, for example (N, t) is quasi-discriminating since both (T_1, t_1) and (T_2, t_2) are discriminating and, additionally, the roots of these trees are labeled differently from the root of (N, t) (ensured by condition (iii)(b)).

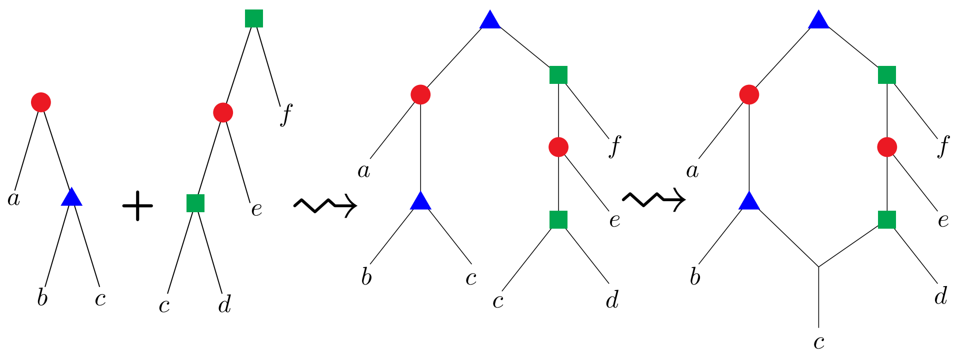
Conversely, say an elementary galled tree (N, t) such as the one in Figure 23 is given. It explains some edge-colored graph Σ (namely, Σ in Figure 22) that can be shown to be a polar-cat. The subgraphs Ω_1 and Ω_2 are defined as the subgraphs of Σ induced by the leaf-children of the left and right sides of the underlying cycle of N, respectively. The caterpillar trees that explain Ω_1 and Ω_2 can then be constructed by the reverse procedure of the previous paragraph, and the join-color will be the label of the root of N.

The ideas presented above will be formalized in Lemma 3.3, Definition 3.4 and Lemma 3.5. To prove that an edge-colored graph is a polar-cat if and only if it is primitive and explained by a galled tree, we will need to focus on how modules are and are not exhibited in galled trees. This is done in Lemma 3.6 and Lemma 3.7.

3.2 Proof of Theorem 3.2 and related results

In this section we introduce and prove a handful of lemmas, culminating in the proof of Theorem 3.2. We begin with the following.

Lemma 3.3. *If (N, t) is a strong elementary and quasi-discriminating galled tree, then $\mathcal{G}(N, t)$ is a polar-cat. In particular, $\mathcal{G}(N, t)$ is a (v, Ω_1, Ω_2)-polar-cat, where v is the child of the hybrid vertex of N, and the subgraphs Ω_1 and Ω_2 are the the induced subgraphs on the leaves belonging to the two respective sides of the underlying cycle of N.*



(T_1, t_1) (T_2, t_2) (N, t)

Figure 23: A visualization of the principal idea for how the structure of polar-cats imply they may be explained by an elementary, quasi-discriminating tree and vice versa. See Example 3.2 for details.

Proof. Let (N, t) be a strong, elementary and quasi-discriminating galled tree and let $\Sigma = (V, \sigma)$ denote the edge-colored graph it explains, i.e. $\Sigma = \mathcal{G}(N, t)$. Since N is strong and elementary, its inner vertices are contained in two directed paths P^1 and P^2 that meet only in the root ρ and the hybrid vertex η. Let W_i consist of the leaf-children of the vertices of P^i and put $\Omega_i = \Sigma[W_i]$, for $i \in \{1, 2\}$. Furthermore, let v be the child of η. Lastly, let ● be the color $t(\rho)$. Note that by construction and N being strong, Ω_1 and Ω_2 immediately satisfies condition (i) of Definition 3.1.

Now, Lemma 2.1 ensures that $t(\mathrm{lca}(x, y)) = t(\rho) = \bullet$ for each $x \in W_1 \setminus \{v\}$ and $y \in W_2 \setminus \{v\}$, since inner vertices of P^1 are incomparable with inner vertices of P^2. Since $\sigma(xy) = t(\mathrm{lca}(x, y))$ for all distinct $x, y \in V$ we thus have $\Sigma - v = (\Omega_1 - v) \otimes_{\bullet} (\Omega_2 - v)$. Thus condition (ii) is satisfied.

We now construct a caterpillar tree that explains Ω_1; analogous arguments can be made for Ω_2. Suppose u is the child of ρ that lies on P^1, and let (T_1, t_1) be the subnetwork $N(u)$. Clearly, T_1 is a caterpillar tree, and only η_N is suppressed while constructing $N(u)$. Moreover, if we let u' denote the lowest (w.r.t. $\preceq_N$) inner vertex of P^1, then the leaf v must, by construction, be a child of u' in T_1 and hence part of its cherry. By construction, the leaf set of T_1 is precisely W_1, the vertices of Ω_1. We now show that (T_1, t_1) explains Ω_1. Consider any distinct $x, x' \in W_1$. Since T_1 is a caterpillar, we have

$$\mathrm{lca}_{T_1}(x, x') = \max_{\succeq_{T_1}} \{\mathrm{parent}_{T_1}(x), \mathrm{parent}_{T_1}(x')\}.$$

On the other hand, we similarly have that

$$\mathrm{lca}_N(x, x') = \max_{\succeq_N} \{\mathrm{parent}_N(x), \mathrm{parent}_N(x')\},$$

so $\mathrm{lca}_N(x, x') = \mathrm{lca}_{T_1}(x, x')$, for all x and x'. Since t_1 is the restriction of t, this means that (T_1, t_1) indeed explains Ω_1.

Now, (T_1, t_1) is discriminating since (N, t) is quasi-discriminating, and the suppressed vertex was a hybrid vertex in N. The quasi-discriminating nature of N in particular implies that $t_1(\rho_{T_1}) \neq \bullet$. This concludes that Σ is a polar-cat. $\square$

Given a galled tree (N, t), the operation of constructing the edge-colored graph $\mathcal{G}(N, t)$ has no inherent difficulties – instead the proof of Lemma 3.3 involved showing certain properties of $\mathcal{G}(N, t)$ (rather than proving its existence) for a particular type of galled tree. Proving the existence of a galled tree that explains a given edge-colored graph is, in general, a trickier endeavor. However, the restricted structure of polar-cats makes it possible to explicitly construct a galled tree which explains it. The main idea is to connect the caterpillars that explain the subgraphs Ω_1 and Ω_2 of a (v, Ω_1, Ω_2)-polar-cat with an additional root-vertex and by merging their respective occurrences of the leaf v under a common hybrid vertex. This procedure is formalized as follows (c.f. [19, Def. 4.14]).

Definition 3.4. Let Σ be a (v, Ω_1, Ω_2)-polar-cat with join-color k. Let (T_1, t_1) and (T_2, t_2) be the discriminating caterpillar tree that explains Ω_1 respectively Ω_2. The directed graph $N := \mathcal{N}(v, \Omega_1, \Omega_2)$ is constructed as follows.

1. We may, without loss of generality, assume T_1 and T_2 are vertex-disjoint except for the leaf v. To distinguish the occurrences of v in the two trees, exchange v in T_i with v_i for $i \in \{1, 2\}$. Then begin constructing N by taking the disjoint union of (the modification of) T_1 and T_2.

2. Add a new root vertex ρ_N to N, along with the edges (ρ_N, ρ_{T_1}) and (ρ_N, ρ_{T_2}).

3. Identify the leaves v_1 and v_2 of N into a new hybrid vertex denoted η_N, add a new occurrence of the leaf v and the edge (η_N, v).

A labeling $t := t(v, \Omega_1, \Omega_2)$ of $\mathcal{N}(v, \Omega_1, \Omega_2)$ is defined by

$$t(u) := \begin{cases} t_1(u) & \text{if } u \in V(T_1) \\ t_2(u) & \text{if } u \in V(T_2) \\ k & \text{if } u \in \{\rho_N, \eta_N\} \end{cases}$$

for each $u \in V(\mathcal{N}(v, \Omega_1, \Omega_2))$. The value $t(\eta_N)$ may be arbitrarily defined: for definiteness we may choose $t(\eta_N) = k$, as done above.

For a visualization of the definition, refer back to Figure 23. Since the trees T_1 and T_2 used in the construction of $N = \mathcal{N}(v, \Omega_1, \Omega_2)$ are caterpillars, it is easily verified that N is a galled tree consisting of a single cycle rooted at ρ_N and terminating at η_N. Note that η_N has a single child, namely, the leaf v, ρ_N has no leaf-children and all other inner vertices have precisely one leaf-child. In other words, N is an elementary galled tree. The labeling $t = t(v, \Omega_1, \Omega_2)$ is well-defined since v is the only vertex shared by T_1 and T_2, and since v is a leaf in both trees we have $t_1(v) = t_2(v)$.

Moreover, the labeled galled tree (N, t) is quasi-discriminating since both T_1 and T_2 are discriminating, since polar-cats are polarizing (i.e. the root of N is distinctly labeled from its children) and since the only newly introduced inner vertex except the root is the hybrid vertex η_N.

Next, recall that both Ω_1 and Ω_2 have at least two vertices each — corresponding to at least two leaves of T_1 and T_2 — so the cycle of N cannot be an unbalanced weak cycle. If, additionally, the polar-cat has at least four vertices then we can conclude that at least one of Ω_1 and Ω_2 has at least three vertices, implying that T_1 and T_2 has three or more leafs, so that N by construction is not a balanced weak cycle.

Lastly, note that per construction we have $(T_1, t_2) = N(u)$ and $(T_2, t_2) = N(v)$, where u and v are the two children of ρ_N. We collect these observations about Definition 3.4 in the following.

Observation 3.2. If Σ is a (v, Ω_1, Ω_2)-polar-cat, then the directed graph $\mathcal{N}(v, \Omega_1, \Omega_2)$ is a galled tree, and $t(v, \Omega_1, \Omega_2)$ is a well-defined labeling of $\mathcal{N}(v, \Omega_1, \Omega_2)$. In particular, $(\mathcal{N}(v, \Omega_1, \Omega_2), t(v, \Omega_1, \Omega_2))$ is elementary and quasi-discriminating. If, additionally, Σ has at least four vertices, then $(\mathcal{N}(v, \Omega_1, \Omega_2), t(v, \Omega_1, \Omega_2))$ is strong. The subgraphs Ω_1 and Ω_2 are explained by the caterpillar trees $N(u)$ and $N(v)$, where u and v are the children of the root of $\mathcal{N}(v, \Omega_1, \Omega_2)$.

More importantly, we have the following result.

Lemma 3.5. *Every (v, Ω_1, Ω_2)-polar-cat is explained by the galled tree $(\mathcal{N}(v, \Omega_1, \Omega_2), t(v, \Omega_1, \Omega_2))$.*

Proof. Let $\Sigma = (V, \sigma)$ be a (v, Ω_1, Ω_2)-polar-cat and let (N, t) denote the elementary and quasi-discriminating galled tree $(\mathcal{N}(v, \Omega_1, \Omega_2), t(v, \Omega_1, \Omega_2))$. Let u_1 and u_2 be the two children of ρ_N, so that $N(u_1)$ and $N(u_2)$ explain Ω_1 and Ω_2, respectively (see Observation 3.2). Let W_1 and W_2 denote the vertices of Ω_1 respective Ω_2.

Throughout this proof, we say that N explains the edge $\{x, y\}$ if $\sigma(xy) = t(\mathrm{lca}_N(x, y))$ for the distinct vertices $x, y \in V$. To show that (N, t) explains Σ we need to show that N explains every edge of Σ. Consider any two leaves x and x' of N. By Lemma 2.1 we have $t(\mathrm{lca}_N(x, x')) = t(\rho_N)$ if the parents of x and x' are incomparable in N. This happens if and only if, say, x is a leaf of $N(u_1)$, x' is a leaf of $N(u_2)$ and $v \notin \{x, x'\}$. But then $x \in W_1 \setminus \{v\}$ and $x' \in W_2 \setminus \{v\}$, which implies $\sigma(xx')$ equals the join-color k of Σ. By construction of t, we indeed have $t(\rho_N) = k$, so that N correctly explains the edge $\{x, x'\}$.

If, on the other hand, the parents of x and x' are comparable in N, then x and x' are both leaves of $N(u_i)$, for $i \in \{1, 2\}$. We may without loss of generality assume that $\text{parent}_N(x) \succeq_N \text{parent}_N(x')$, so that Lemma 2.1 implies $t(\text{lca}_N(x, x')) = t(\text{parent}_N(x))$. Since on the inner vertices of $N(u_i)$ the labeling t restricts to t_i we thus have

$$t(\text{lca}_N(x, x')) = t_i(\text{parent}_{N(u_i)}(x)) = \sigma(xx').$$

That is, N correctly explains the edge $\{x, x'\}$ in this case as well. $\square$

We now investigate the structures of modules in polar-cats, and prove the following.

Lemma 3.6. *If Σ is a polar-cat, then Σ is primitive.*

Proof. Suppose $\Sigma = (V, \sigma)$ is a polar-cat. As seen in Observation 3.1, Σ must be a rainbow triangle if it has three or less vertices, and rainbow triangles are primitive edge-colored graph (c.f. Observation 2.2). We may thus assume $|V| \geq 4$. The combination of Observation 3.2 and Lemma 3.5 ensures the existence of a strong elementary quasi-discriminating galled tree (N, t) that explains Σ. We let C denote the unique underlying cycle of N and let ρ and η denote its root and unique hybrid vertex, respectively. In what follows, we let x^C, y^C, u^C etc. denote the unique parent on C of the leaves x, y, and u etc.

Let $M \subseteq V$ be a set of vertices such that $2 \leq |M| < |V|$. We will show that M is not a module by finding distinct $x, y \in M$ and some $u \notin M$ such that $\sigma(xu) \neq \sigma(uy)$. We will use Lemma 2.1 repeatedly, without explicit reference.

Note that since $\succ_N$ is a partial order and each leaf has a distinct parent, we always have that $y^C \succ_N x^C$, $x^C \succ_N y^C$ or $x^C \parallel_N y^C$ for any two distinct $x, y \in V$. Moreover, no leaf has ρ as a parent. These facts will make it possible to distinguish three possible cases; we show the first, and then motivate and prove the remaining two.

Case 1: There exist some $x, y \in M$ and some $u \notin M$ such that $x^C \succ_N u^C \succ_N y^C$

We may, without loss of generality, assume that $x, y \in M$ are taken so that there is no $z \in M$ with $x^C \succ_N z^C \succ_N y^C$. In particular the case assumption then implies that the (unique) non-leaf child u^C of x^C is parent to a leaf u such that $u \notin M$. The quasi-discriminating nature of N ensures $t(x^C) \neq t(u^C)$. We thus have

$$\sigma(xu) = t(x^C) \neq t(u^C) = \sigma(uy),$$

and M is not a module.

With Case 1 in mind, we may in the remaining cases assume that for each pair of vertices $x, y \in M$, it holds that $u \in M$ for any leaf u such that $x^C \succeq_N u^C \succeq_N y^C$. Since there cannot be vertices $x, y \in M$ and $u \in M$ with $x^C \succ_N u^C \succ_N y^C$ we may distinguish between the following two cases. Either we have that $u^C \succ_N x^C$ for every $x \in M$ and $u \notin M$ such that x^C and u^C are comparable, or we have that $x^C \succ_N u^C$ for every $x \in M$ and $u \notin M$ such that x^C and u^C are comparable. We now prove these two cases (recall that $u \parallel_N v$ denotes u and v being incomparable w.r.t. $\succ_N$). For the reader's convenience, we provide a visual overview of the cases in Figure 24 and Figure 25.

Case 2: $u^C \succ_N x^C$ or $u^C \parallel_N x^C$ for each $x \in M$ and each $u \notin M$. In particular we must have that the unique child of η, here denoted z, is an element of M, since $v \succeq_N \eta$ for every internal vertex v of N. Moreover, since $|M| \geq 2$ we also know there is some $y \in M$ such that y^C is a parent of η. If there is some $u \notin M$ with $u^C \parallel_N y^C$, then there in particular must exist some $v \notin M$ where $v^C \parallel_N y^C$ and v^C is a child of ρ. By N being quasi-discriminating we, in this case, have that

$$\sigma(yv) = t(\rho) \neq t(v^C) = \sigma(vz).$$

If, on the other hand, each $u \notin M$ has a parent such that $u^C \succ_N y^C$, then there is a vertex $v \notin M$ which has a parent v^C that is a child of ρ. Moreover, since C is a strong cycle there is some $w \in M$ with $w^C \parallel_N v^C$, and we have

$$\sigma(wv) = t(\rho) \neq t(v^C) = \sigma(vy),$$

again using that N is quasi-discriminating. Either way, M cannot be a module.

Case 3: $x^C \succ_N u^C$ or $x^C \parallel_N u^C$ for each $x \in M$ and each $u \notin M$. First assume that there exist $x, y \in M$ such that $t(x^C) \neq t(y^C)$. Since, in particular, the child u of η does not lie in M (if so, then the case assumption would imply $M = V$), we have

$$\sigma(xu) = t(x^C) \neq t(y^C) = \sigma(uy).$$

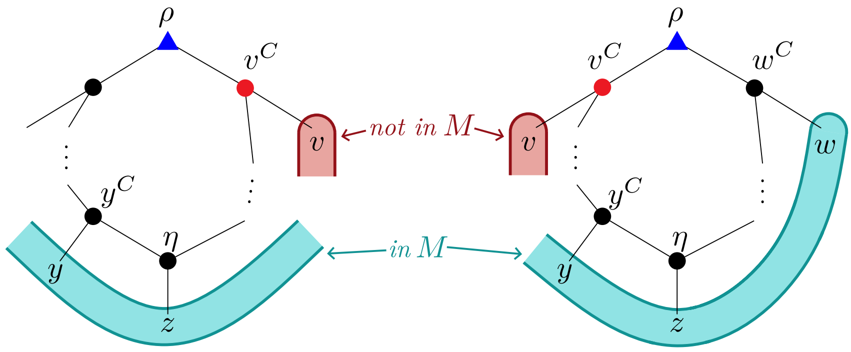
If, on the other hand, $t(x^C) = t(y^C)$ for each $x, y \in M$ then the case assumption and N being quasi-discriminating together implies both that $M = \{x, y\}$ and that x^C and y^C are the two distinct children of ρ. In particular $t(y^C) \neq t(\rho)$. Since C is a strong cycle, this further means that there exist some $v \notin M$ where $v^C \neq \eta$. We may, without loss of generality, suppose $y^C \succ_N v^C$, so that

$$\sigma(xv) = t(\rho) \neq t(y^C) = \sigma(vy).$$

Again, it is not possible that M is a module.

We have thus shown that no $M \subseteq V$ such that $2 \leq |M| < |V|$ is a module. Hence $\mathbb{M}(\Sigma)$ consists of trivial modules only, i.e. Σ is primitive.

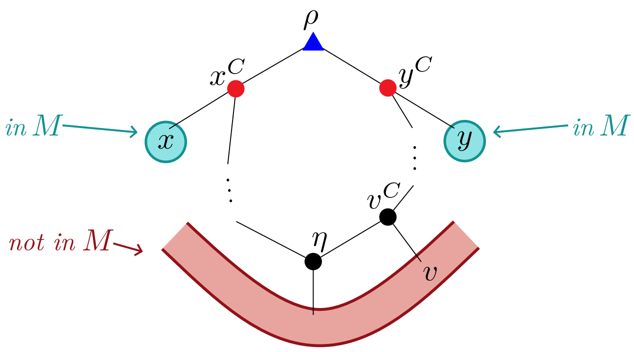
□



$$\exists u \notin M : u^C ||_N y^C \qquad \nexists u \notin M : u^C ||_N y^C$$

$$\Rightarrow \sigma(yv) = t(\rho) \neq t(v^C) = \sigma(vz) \qquad \Rightarrow \sigma(wv) = t(\rho) \neq t(v^C) = \sigma(vy)$$

Figure 24: A visual aid for Case 2 in the proof of Lemma 3.6. The first subcase is depicted on the left, and the second subcase on the right. The labels ▲ and ● are distinct, while the labels • are unknown.



$$\forall x, y \in M : t(x^C) = t(y^C)$$

$$\Rightarrow \exists v \notin M : \sigma(vx) = t(\rho) \neq t(y^C) = \sigma(yv)$$

Figure 25: A visual aid for Case 3 in the proof of Lemma 3.6. Only the second subcase is depicted, since the first involve less notation. The labels ▲ and ● are distinct, while the labels • are unknown.

Before we continue the investigation of edge-colored graphs that are explained by galled trees we make a note about modules in galled trees.

Observation 3.3. Let (N, t) be a galled tree, and assume C is a cycle of N with root ρ. The leaves of the subnetwork $N(\rho)$ is a module of $\mathcal{G}(N, t)$, since for every $x \in \mathcal{L}(N(\rho_C))$ and every $y \notin \mathcal{L}(N(\rho_C))$, we have that $\mathrm{lca}_N(x, y) = \mathrm{lca}_N(\rho_C, y)$.

For primitive edge-colored graphs we also have the following result.

Lemma 3.7. *Let Σ be a primitive edge-colored graph with at least four vertices. If Σ is explained by a galled tree (N, t), then (N, t) is strong, elementary and quasi-discriminating.*

Proof. Suppose (N, t) is a galled tree that explains the primitive edge-colored graph $\Sigma = (V, \sigma)$, and assume that $|V| \geq 4$. The primitivity of Σ in particular means that Σ is not $\mathfrak{unp}$, hence Theorem 2.5 ensures that N has at least one cycle C. Suppose that the root ρ_C of C is not the same as the root ρ_N. Then the set $\mathcal{L}(N(\rho_C))$, which has at least two vertices, is a proper subset of V. By Observation 3.3 it is thus a nontrivial module of Σ, which is a contradiction to the primitivity of Σ. Since C was arbitrarily chosen, we thus conclude that N has a single cycle C such that $\rho_C = \rho_N$.

Let $V^0(C)$ denote the (inner) vertices of the cycle C. For each $v \in V^0(C)$ we introduce the set

$$\mathcal{L}^C(v) := \left\{x \in \mathcal{L}(N(v)) \,|\, \forall u \in V^0(C) \text{ s.t. } v \succeq_N u \text{ we have } x \not\preceq_N u\right\}.$$

In other words, $\mathcal{L}^C(v)$ consists of the leaves whose ancestors on C consist of v and ancestors of v. To show that N is elementary, it suffices to show that $\mathcal{L}^C(\rho_N) = \emptyset$ and that $|\mathcal{L}^C(v)| = 1$ for each $v \in V^0(C) \setminus \{\rho_N\}$. First note that for any $v \in V^0(C)$, any $x, x' \in \mathcal{L}^C(v)$ and any $y \in \mathcal{L}(N) \setminus \mathcal{L}^C(v)$ we have $\mathrm{lca}_N(x, y) = \mathrm{lca}_N(x', y)$. This means that $\mathcal{L}^C(v)$ is either empty or a module of Σ. For $v \in V^0(C)$ distinct from ρ_N, v must have at least some descendent not on C (otherwise it has both in- and outdegree one), thus $\mathcal{L}^C(v)$ is not empty. Clearly, $\mathcal{L}^C(v)$ is a proper subset of V, so the triviality of the modules of Σ implies $|L^C(v)| = 1$. On the other hand, if $L^C(\rho_N) \neq \emptyset$ then for each $x \in L^C(\rho_N)$ and all distinct leaves y, y' not in $L^C(\rho_N)$, we have $\mathrm{lca}(x, y) = \rho_N = \mathrm{lca}(x, y')$, making $\mathcal{L}(N) \setminus L^C(\rho_N)$ into a non-trivial module. This contradicts primitivity of Σ, and we may conclude that N is an elementary galled tree.

We now show that N is a strong galled tree, i.e. that C is a strong cycle. Since N is elementary and has at least four leafs, C cannot be a balanced weak cycle; the latter has three leaves only. If C was an unbalanced weak cycle, then Lemma 2.3 would imply the existence of a tree (T', t') such that $\mathcal{G}(T', t') = \Sigma$. But if Σ is explained by a tree, then it is $\mathfrak{unp}$ by Theorem 2.5. That contradicts Σ being primitive, hence X is not an unbalanced weak cycle either.

Lastly, suppose for contradiction that (N, t) is not quasi-discriminating and let $v, v' \in V^0(C)$ be adjacent tree vertices of C such that $t(v) = t(v')$. In particular $\eta_C \notin \{v, v'\}$ since η_C, by definition, is a hybrid vertex and thus not a tree vertex. Let u denote the unique leaf-child of η_C. Without loss of generality, we may suppose $\rho_N \succeq_N v \succ v' \succ \eta_C$. Denote the unique leaf-child of v' with x'. First assume $\rho_N = v$, so that v has no leaf-children. Moreover, this assumption enforces that there is no leaf $y \in V \setminus \{x'\}$ such that $\mathrm{parent}_N(y) \succ_N v'$. Hence $v' \succ_N \mathrm{parent}_N(y)$ or $v' \parallel_N \mathrm{parent}_N(y)$ for all $y \in V \setminus \{x'\}$, from which Lemma 2.1 implies that $\mathrm{lca}_N(x', y) \in \{\rho_N, v'\} = \{v, v'\}$ for all $y \in V \setminus \{x'\}$. Hence $t(v') = t(v)$ implies that $V \setminus \{x'\}$ is a (necessarily non-trivial) module of Σ, which contradicts the primitivity of Σ.

Now assume that $\rho_N \neq v$, so that v also has a unique leaf-child x. Since v and v' are adjacent, the following holds

$$\begin{aligned} \mathrm{parent}_N(y) \succ_N v &\iff \mathrm{parent}_N(y) \succ_N v' \\ v \succ_N \mathrm{parent}_N(y) &\iff v' \succ_N \mathrm{parent}_N(y) \\ \mathrm{parent}_N(y) \parallel_N v &\iff \mathrm{parent}_N(y) \parallel_N v' \end{aligned} \tag{7}$$

for all $y \in V\setminus\{x, x'\}$. Combining (7) with lemma 2.1 implies that $\mathrm{lca}_N(xy) = \mathrm{parent}_N(y) = \mathrm{lca}_N(x'y)$ for all $y \in V \setminus \{x, x'\}$ such that $\mathrm{parent}_N(y) \succ_N v$, while $\mathrm{lca}_N(xy) = \rho_N = \mathrm{lca}_N(x'y)$ for all $y \in V \setminus \{x, x'\}$ such that $\mathrm{parent}_N(y) \parallel_N v$. Lastly, $\mathrm{lca}_N(xy) = v$ respectively $\mathrm{lca}_N(x'y) = v'$ for all $y \in V \setminus \{x, x'\}$ such that $v \succ_N \mathrm{parent}_N(y)$. In all three cases we have $t(\mathrm{lca}_N(xy)) = t(\mathrm{lca}_N(x'y))$ for all $y \in V \setminus \{x, x'\}$. Hence $\{x, x'\}$ is a non-trivial module of Σ, again contradicting its primitivity. This concludes that (N, t) is a strong, elementary and quasi-discriminating galled tree. □

We are now ready to return to the main result of this section: Theorem 3.2.

Proof of Theorem 3.2. By Observation 3.1, a rainbow triangle is the only polar-cat with three vertices (and there are no polar-cats with fewer vertices). Moreover, by Observation 2.2 it is the only primitive edge-colored graph on three vertices, and it is explained by a (quasi-discriminating elementary) galled tree. Thus the three statements are equivalent for edge-colored graphs with three vertices.

From here on we may thus consider graphs on at least four vertices. That every edge-colored graph explained by a strong, elementary and quasi-discriminating galled tree is a polar cat (the implication (a) $\Longrightarrow$ (b)) follows from Lemma 3.3. That polar-cats are explained by galled trees and are primitive (the implication (b) $\Longrightarrow$ (c)) follows from Lemma 3.5 respectively Lemma 3.6. Lastly, that primitive edge-colored graphs explained by a galled trees are, in particular, explained by a strong, elementary and quasi-discriminating galled tree (the implication (c) $\Longrightarrow$ (a)) holds by Lemma 3.7. □

We end this section by returning to one of the first statements we made about edge-colored graphs that are explained by quasi-discriminating galled trees: that they are primitive, yet almost $\mathfrak{unp}$. The precise statement is the following.

Proposition 3.8. *If Σ is a (v, Ω_1, Ω_2)-polar-cat, then $\Sigma - v$ is $\mathfrak{unp}$.*

Proof. Let Σ be a (v, Ω_1, Ω_2)-polar-cat. By definition, Ω_1 and Ω_2 are explained by caterpillar trees (T_1, t_1) and (T_2, t_2), respectively. Construct a tree (T_1', t_1') from (T_1, t_1) by (1) removing the leaf v (which, by definition, is part of its cherry), (2) suppressing the vertex $\mathrm{parent}(v)$ and (3) putting t_1' as the restriction of t_1 to the remaining vertices. Then (T_1', t_1') has only a single vertex, or is a caterpillar tree. Moreover, it is easily verified to explain $\Omega_1 - v$, as the only suppressed vertex $\mathrm{parent}(v)$ satisfies $\mathrm{parent}(v) = \mathrm{lca}_{T_1}(x, y)$ only if $v \in \{x, y\}$. The tree (T_2', t_2') may then be constructed analogously from (T_2, t_2), so that $\mathcal{G}(T_2', t_2') = \Omega_2 - v$.

Now, since $\Sigma - v = (\Omega_1 - v) \otimes_k (\Omega_2 - v)$ for some color k, the trees (T_1', t_1') and (T_2', t_2') may be combined into a tree (T, t) that explains Σ as follows. Take the disjoint union of T_1' and T_2', then add a new root ρ_T and the edges $(\rho_T, \rho_{T_1'})$ and $(\rho_T, \rho_{T_2'})$. Define the labeling t as

$$t(v) := \begin{cases} t_1'(v) & \text{if } v \in V(T_1') \\ t_2'(v) & \text{if } v \in V(T_2') \\ k & \text{otherwise i.e. if } v = \rho_T. \end{cases}$$

for each $v \in V(T)$. Since (T, t) is a tree that explains $\Sigma - v$, Theorem 2.5 implies $\Sigma - v$ is $\mathfrak{unp}$. $\square$

The converse statement of Proposition 3.8 is not true, however. Consider, for example, the edge-colored graph in Figure 26. Clearly, $\Sigma - v$ is $\mathfrak{unp}$, as $\Sigma - v$ has only one edge-color. At the same time, Σ is not a (v, Ω_1, Ω_2)-polar-cat since however we choose Ω_1 and Ω_2 one of them will be a rainbow triangle — for preciseness, the only possible subgraphs satisfying condition (i) of Definition 3.1 are $\{\Omega_1, \Omega_2\} = \{\Sigma[\{v, x\}], \Sigma[\{v, y, z\}]\}$, $\{\Omega_1, \Omega_2\} = \{\Sigma[\{v, y\}], \Sigma[\{v, x, z\}]\}$ or $\{\Omega_1, \Omega_2\} = \{\Sigma[\{v, z\}], \Sigma[\{v, x, y\}]\}$. In fact, there is no other vertex $u \neq v$ such that Σ is a (u, Ω_1, Ω_2)-polar-cat. To see this, simply note that $\Sigma - u$ is a rainbow triangle for each $u \in \{x, y, z\}$, which is not a primitive graph. The contrapositive of Proposition 3.8 thus implies that Σ is not a (u, Ω_1, Ω_2)-polar-cat.

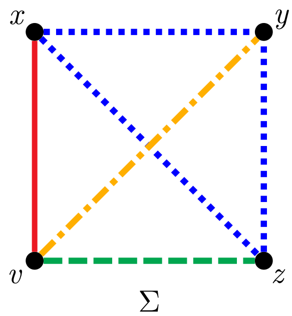


Figure 26: An edge-colored graph Σ for which $\Sigma - v$ is $\mathfrak{unp}$, but Σ is no polar-cat. Σ thus shows that the converse of Proposition 3.8 is, in general, not true.

4 Uniqueness

We now wish to understand when a primitive edge-colored graph is explained by a *distinct* galled tree. To prove any statement about this type of uniqueness, we will need to investigate polar-cats closer. Since much of the structure of a (v, Ω_1, Ω_2)-polar-cats is related to Ω_1 and Ω_2 being explained by caterpillar trees, it is not too surprising that we need to devote some space for results on caterpillar-explainable edge-colored graphs. This is done in Section 4.1. Recall that we say that a polar-cat Σ is explained by a unique galled tree (N, t) if every galled tree (N', t') such that $\mathcal{G}(N', t') = \Sigma$ satisfies $(N, t) \simeq (N', t')$. As we will see, every polar-cat has what we call a fixpoint if and only if it is explained by a unique galled tree. We will then find three sufficient conditions for a primitive edge-colored graph being explained by a unique galled tree. In the last subsection, we investigate what happens when these three sufficient conditions all fail at the same time.

4.1 Structure implied from being caterpillar-explainable

We now state and prove three lemmas in rapid succession.

Lemma 4.1. *Suppose Σ is an edge-colored graph explained by the caterpillar tree (N, t). If $\Sigma = \Sigma[X] \otimes_{\bullet} \Sigma[Y]$ for some disjoint, nonempty subsets $X, Y \subseteq V(\Sigma)$, then $t(\rho_N) = \bullet$.*

Proof. Let (N, t) denote the caterpillar tree that explains the edge-colored graph $\Sigma = (V, \sigma)$. Assume $X, Y \subseteq V$ are nonempty sets of vertices such that $X \cap Y = \emptyset$ and

$$\Sigma = \Sigma[X] \otimes_{\bullet} \Sigma[Y] \tag{8}$$

for some color $\bullet$. By definition we must then have that $X \cup Y = V$, so any vertex of V is an element of either X or Y.

Since N is a caterpillar ρ_N has a leaf-child v. Clearly, $\mathrm{lca}(x,v) = \rho_N$ for each $x \in V \setminus \{v\}$, so $\sigma(vx) = \sigma(vx')$ for all $x, x' \in V \setminus \{v\}$. Now, since $v \in X$ or $v \in Y$, (8) implies $\sigma(vx) = \bullet$ at least for some $x \in V \setminus \{v\}$, from which $t(\rho_N) = \bullet$ follows. $\square$

Lemma 4.2. *If Σ is an edge-colored graph that has at least three vertices and is caterpillar-explainable, then $\Sigma - v$ is caterpillar-explainable.*

Proof. Let (N,t) denote a caterpillar tree that explains the edge-colored graph $\Sigma = (V, \sigma)$. Take any $v \in V$. If v is the leaf-child of ρ_N and u denotes the child of ρ_N such that $u \neq v$, then it is obvious that the subnetwork $N(u)$, which is a caterpillar, explains $\Sigma - v$. On the other hand, if $\mathrm{parent}_N(v) \neq \rho_N$, then the tree (N', t') constructed from (N,t) by (1) deleting v and its incident edge, (2) suppressing the vertex $\mathrm{parent}_N(v)$, and (3) defining t' as the restriction of t to $N(T) \setminus \{v, \mathrm{parent}_N(v)\}$ is a caterpillar tree as well. In particular it explains $\Sigma - v$, since $\mathrm{lca}_N(x,y) = \mathrm{lca}_{N'}(x,y)$ for each pair of distinct $x, y \in V \setminus \{v\}$ follows directly from construction. $\square$

Lemma 4.3. *Let $\Sigma = (V, \sigma)$ be an edge-colored graph with n vertices. Suppose Σ is explained by a discriminating caterpillar tree, and let $x \in V$ be part of its cherry. If u_1, u_2, ..., u_k are vertices such that*

$$\sigma(xu_1) = \sigma(xu_2) = \ldots = \sigma(xu_k),$$

then $k \leq \lfloor n/2 \rfloor$. Additionally, for each edge-color $\bullet$ of Σ there exists some vertex $x \neq y \in V$ such that $\sigma(xy) = \bullet$.

Proof. Let (N,t) be the discriminating caterpillar tree that explains the edge-colored graph $\Sigma = (V, \sigma)$. Note that N has precisely $n-1$ inner vertices, where $n := |V|$. The result is immediately clear when $n \in \{1, 2\}$, so suppose $n \geq 3$. We may denote the inner vertices of N by v_1, v_2, ..., v_{n-1} where

$$v_1 \succ_N v_2 \succ_N \ldots \succ_N v_{n-1}.$$

Furthermore, let u_i be the unique leaf-child of v_i for each $i = 1, 2, \ldots, n-2$, and let x denote one of the leaf-children of u_{n-1}, so that x is part of the cherry of N. Since $t(v_i) \neq t(v_{i+1})$ for each $i = 1, \ldots, n-2$ at most $\lfloor n/2 \rfloor$ inner vertices may have the same label in (N,t). Now, we have

$$\sigma(xu_i) = t(\mathrm{lca}(x, u_i)) = t(v_i)$$

for each $i = 1, \ldots, n-2$, so that means at most $\lfloor n/2 \rfloor$ vertices have an edge to x colored with the same color.

For the second statement, simply note that for each v_i with $i = 1, \ldots, n-1$ there some leaf u, namely $u = u_i$, such that $v_i = \mathrm{lca}(ux)$. For each edge-color $\bullet$ of Σ, there is some $i = 1, \ldots, k$ such that $t(v_i) = \bullet$, from which the statement follows. $\square$

4.2 Sufficient conditions for uniqueness

Before properly investigating uniqueness-results, we study rainbow triangles in the following proposition. It should not be surprising that rainbow triangles are of particular interest — for one thing it appears in the statement of Theorem 3.2. Moreover, by Theorem 2.5, the existence of a rainbow triangle in an edge-colored graph Σ implies that Σ is not $\mathfrak{unp}$. One can thus think of rainbow triangles as a road block removing the possibility of edge-colored graphs being explained by trees. At the same time, rainbow triangles can be explained by galled trees (see Observation 2.2).

Proposition 4.4. *Let Σ be a polar-cat. If Σ has at least three edge-colors, then Σ contains a rainbow triangle. Moreover, if x, y and z are three vertices such that $\Sigma[\{x, y, z\}]$ is a rainbow triangle then Σ is, in particular, a (v, Ω_1, Ω_2)-polar-cat, for at least some $v \in \{x, y, z\}$, while $u \in V(\Omega_1)$ and $w \in V(\Omega_2)$ for $\{u, w\} = \{x, y, z\} \setminus \{v\}$.*

Proof. Let $\Sigma = (V, \sigma)$ be a polar-cat with at least three edge-colors. If Σ itself is not a rainbow triangle, Theorem 3.2 ensures the existence of a strong, elementary, quasi-discriminating galled tree (N, t) that explains Σ. Let x^C and y^C be the children of ρ_N, and denote their respective leaf-children by x and y. Moreover, let v be the unique leaf-child of the hybrid η_N. Since N is strong, we know that $\eta_N \notin \{x^C, y^C\}$. Also recall that the color $t(\rho_N)$ is distinct from $t(x^C)$ and $t(y^C)$.

Consider first the case when $t(x^C) \neq t(y^C)$. Since Lemma 2.1 ensures that $\operatorname{lca}(x, y) = \rho$, $\operatorname{lca}(x, v) = x^C$ and $\operatorname{lca}(y, v) = y^C$ we have

$$|\{\sigma(xy), \sigma(xv), \sigma(vy)\}| = |\{t(\rho), t(x^C), t(y^C)\}| = 3,$$

i.e. the three vertices x, y and v induce a rainbow triangle in Σ.

If, on the other hand, $t(x^C) = t(y^C)$, then there must exist a leaf $z \notin \{x, y, v\}$ of N such that its parent z^C is labeled with some color different from $t(\rho)$ and $t(x^C)$ (i.e. different from $t(y^C)$) — otherwise Σ has only two edge-colors. We may, without loss of generality, assume $z^C \prec_N y^C$. Lemma 2.1 again ensures that $\operatorname{lca}(x, z) = \rho$, $\operatorname{lca}(x, v) = x^C$ and $\operatorname{lca}(z, v) = z^C$, so that x, z and v induce a rainbow triangle in Σ.

For the second statement, suppose Σ is a (v, Ω_1, Ω_2)-polar-cat and that x, y, and z are vertices which induce a rainbow triangle in Σ. Note that if $x, y, z \in V(\Omega_i)$ for either $i = 1$ or $i = 2$, then Ω_i contains a rainbow triangle and is, by Theorem 2.5 not $\mathfrak{unp}$ and thus, in particular, not caterpillar-explainable. But Σ is a (v, Ω_1, Ω_2)-polar-cat, so this is impossible. Thus assume that $x, y \in V(\Omega_i)$ and $z \notin V(\Omega_i)$ for $i = 1$ or $i = 2$. In particular this means $z \neq v$. For contradiction we assume $v \notin \{x, y\}$. But then $\sigma(xz) = \bullet = \sigma(yz)$ since $\Sigma - v = (\Omega_1 - v) \otimes_{\bullet} (\Omega_2 - v)$ for some color $\bullet$. This contradicts $\Sigma[\{x, y, z\}]$ being a rainbow triangle. Hence $v \in \{x, y\}$ must hold, and the result follows. $\square$

We are now ready to prove the first type of uniqueness of polar-cats: if we allow ourself to fix the vertex v, then a polar-cat can only be a (v, Ω_1, Ω_2)-polar-cat for particular Ω_1 and Ω_2. The precise statement is as follows.

Proposition 4.5. *If Σ is both a (v, Ω_1, Ω_2)-polar-cat and a (v, Π_1, Π_2)-polar-cat for some fixed vertex v, then $\{\Omega_1, \Omega_2\} = \{\Pi_1, \Pi_2\}$.*

Proof. Let $\Sigma = (V, \sigma)$ be a (v, Ω_1, Ω_2)-polar-cat that is simultaneously a (v, Π_1, Π_2)-polar-cat. By definition of polar-cats, we know that

$$\Sigma - v = (\Omega_1 - v) \otimes_{\textcolor{red}{\bullet}} (\Omega_2 - v) = (\Pi_1 - v) \otimes_{\textcolor{green}{\blacksquare}} (\Pi_2 - v), \tag{9}$$

for some colors $\textcolor{red}{\bullet}$ and $\textcolor{green}{\blacksquare}$. By applying Lemma 2.4 on $\Sigma - v$ we can conclude that $\textcolor{green}{\blacksquare} = \textcolor{red}{\bullet}$, since one way of re-phrasing (9) is that the trivial module $V(\Sigma - v)$ is both a $\textcolor{red}{\bullet}$-series and a $\textcolor{green}{\blacksquare}$-series module of $\Sigma - v$. Also recall that by definition, there exists a discriminating caterpillar-tree (T_X, t_X) that explains X for each $X \in \{\Omega_1, \Omega_2, \Pi_1, \Pi_2\}$. We will show that if $\{\Omega_1, \Omega_2\} \neq \{\Pi_1, \Pi_2\}$, then $t_X(\rho_{T_X}) = \textcolor{red}{\bullet}$ for some $X \in \{\Omega_1, \Omega_2, \Pi_1, \Pi_2\}$, which contradicts condition (ii)(b) of the definition of a polar-cat (Definition 3.1).

For contradiction, assume $\{\Omega_1, \Omega_2\} \neq \{\Pi_1, \Pi_2\}$. Note that $v \in X$ for each $X \in \{\Omega_1, \Omega_2, \Pi_1, \Pi_2\}$, so no two of the four subgraphs have an empty intersection. Moreover, since $V(\Omega_1) \cup V(\Omega_2) = V(\Pi_1) \cup V(\Pi_2) = V$ at least one of Ω_1 and Ω_2 will contain vertices x, y such that $v \notin \{x, y\}$, $x \in V(\Pi_1)$ and $y \in V(\Pi_2)$. We may thus assume, without loss of generality, that neither $V(\Omega_1) \setminus V(\Pi_1)$ nor $V(\Omega_1) \setminus V(\Pi_2)$ are empty.

Now, since (9) in particular implies $\sigma(xy) = \textcolor{red}{\bullet}$ for each $x \in V(\Pi_1) \setminus \{v\}$ and $y \in V(\Pi_2) \setminus \{v\}$ we have

$$\Omega_1 - v = \Omega_1[V(\Omega_1) \setminus V(\Pi_1)] \otimes_{\textcolor{red}{\bullet}} \Omega_1[V(\Omega_1) \setminus V(\Pi_2)]. \tag{10}$$

By assumption Ω_1 has to, in addition to the vertex v, contain at least one element of $V(\Pi_1) \setminus V(\Pi_2)$ and at least one element of $V(\Pi_2) \setminus V(\Pi_1)$, hence $|V(\Omega_1)| \geq 3$. Thus Lemma 4.2 implies $\Omega_1 - v$ is caterpillar-explainable, from which (10) and Lemma 4.1 implies that $t_{\Omega_1}(\rho_{T_{\Omega_1}}) = \textcolor{red}{\bullet}$. As explained, this contradicts (9). Thus $\{\Omega_1, \Omega_2\} = \{\Pi_1, \Pi_2\}$ follows. $\square$

In light of Proposition 4.5, the pursuit of understanding when Σ is a (v, Ω_1, Ω_2)-polar-cat but not a (w, Π_1, Π_2)-polar-cat for $w \neq v$ reduces to understanding when the choice of the vertex v is unique. Formally, we say that a polar-cat Σ *has a fixpoint* if there is a unique vertex v such that Σ is a (v, Ω_1, Ω_2)-polar-cat. Although not explicitly stated, we have in fact already come across a polar-cat without a fixpoint: rainbow triangles.

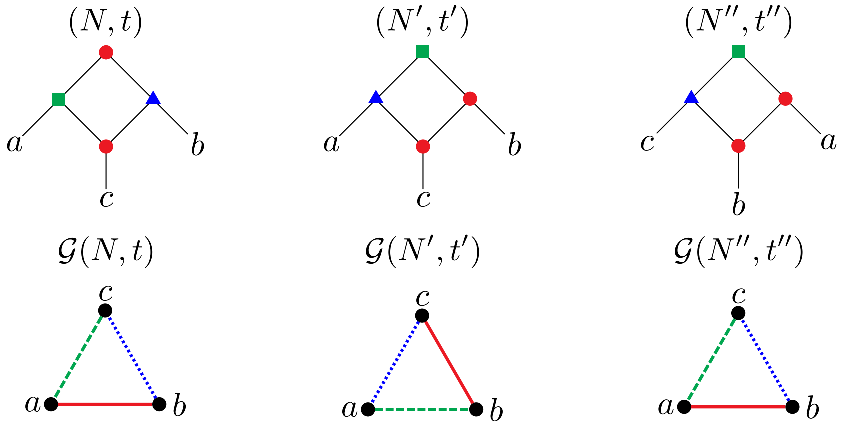


Figure 27: Three galled trees (N,t), (N',t') and (N'',t''), depicted above the respective polar-cat they explain. In particular $(N,t) \simeq (N',t')$ while $(N,t) \not\simeq (N'',t'')$. However, $\mathcal{G}(N,t) \neq \mathcal{G}(N',t')$ while $\mathcal{G}(N,t) = \mathcal{G}(N'',t'')$.

Example 4.1. Rainbow triangles do not have a fixpoint, as the rainbow triangle $\Sigma := \mathcal{G}(N,t)$ of Figure 27 is both a $(c, \Sigma[\{c,a\}], \Sigma[\{c,a\}])$-polar-cat and a $(b, \Sigma[\{b,c\}], \Sigma[\{b,a\}])$-polar-cat. On a related note, we may consider the three galled trees in Figure 27. It is easy to verify that $(N,t) \simeq (N',t')$, while $(N,t) \not\simeq (N'',t'')$: c.f. Observation 2.1. Note also that $\mathcal{G}(N,t) \neq \mathcal{G}(N',t')$ (although, $\mathcal{G}(N,t) \simeq \mathcal{G}(N',t')$, as they are both rainbow triangles). Moreover, $\mathcal{G}(N,t) = \mathcal{G}(N'',t'')$ although the two galled trees (N,t) and (N'',t'') are not isomorphic. In particular this means that there are distinct galled trees that explain a rainbow triangle.

As we will now see, it is no coincident that rainbow triangles are explained by multiple, distinct galled trees and have no fixed point.

Theorem 4.6. *Suppose Σ is a polar-cat. Σ has a fixpoint if and only if there is a unique galled tree (N,t) that explains Σ. Additionally, if Σ has a fixpoint, then the labeling map t of (N,t) is uniquely determined, up to the label of the hybrid vertex of N.*

Proof. First assume that Σ has no fixpoint, i.e. assume that it is both a (v, Ω_1, Ω_2)-polar-cat and a (w, Π_1, Π_2)-polar-cat for vertices $v \neq w$. Then the galled trees

$$(N,t) := (\mathcal{N}(v,\Omega_1,\Omega_2), t(v,\Omega_1,\Omega_2)) \text{ and}$$
$$(N',t') := (\mathcal{N}(w,\Pi_1,\Pi_2), t(w,\Pi_1,\Pi_2))$$

will, by Lemma 3.5, explain Σ. By Observation 3.2, the leaf-child of the respective hybrid-vertex in N and N' is v and w, respectively, so no isomorphism of N and N' as directed graph restricts to the identity map on the leaves. Hence $(N,t) \not\simeq (N',t')$, so that Σ is explained by (at least) two distinct galled trees.

It remains to show the "only if"-direction. To this end, suppose that the polar-cat $\Sigma = (V, \sigma)$ has a fixpoint. In particular, Σ is not a rainbow triangle, as rainbow triangles has no fixpoint (see Ex. 4.1). Suppose that (N, t) and (N', t') are galled trees that explain Σ. We will show that (N, t) and (N', t') are isomorphic. By Theorem 3.2 and Lemma 3.7 both (N, t) and (N', t') are necessarily strong, elementary and quasi-discriminating galled trees. Let v denote the unique leaf-child of the hybrid of N, and put Ω_1 respectively Ω_2 as the induced subgraphs of the leaves that belong to the respective sides of the underlying cycle of N. Introduce v', Ω'_1 and Ω'_2 from N' analogously. Now Lemma 3.3 implies that Σ is a (v, Ω_1, Ω_2)-polar-cat and a $(v', \Omega'_1, \Omega'_2)$-polar-cat. Thus $v = v'$ follows, since Σ has a fixpoint. Moreover, $\{\Omega_1, \Omega_2\} = \{\Omega'_1, \Omega'_2\}$ follows from Proposition 4.5. Without loss of generality, suppose $\Omega_1 = \Omega'_1$ and $\Omega_2 = \Omega'_2$.

We have thus shown that (N, t) and (N', t') are elementary galled trees whose respective sides have the same respective leafs. It remains, loosely speaking, to show that the position of the leaves are the same on the respective side of N and N'. To formalize this, we first need to introduce some notation.

Consider the side $P = u_0u_1u_2 \dots u_k$ of the underlying cycle of N such that the leaves that belong to P are precisely those vertices that are contained in Ω_1. In particular, $u_0 = \rho_N$ and $u_k = \eta_N = \mathrm{parent}_N(v)$. Denote the (unique) leaf-child of u_i by x_i for each $1 \leq i \leq k$ so that, in particular, $x_k = v$ and $V(\Omega_1) = \{x_1, \dots, x_k\}$. Similarly, let $P' = u'_0u'_1 \dots u'_k$ be the side of of the underlying cycle of N' such that the leaves that belong to P' are precisely $\{x_1, \dots, x_k\}$. Again, we necessarily have $u'_0 = \rho_{N'}$ and $u'_k = \eta_N = \mathrm{parent}_{N'}(v)$. By definition we have $\mathrm{parent}_N(x_i) = u_i$ for each $1 \leq i \leq k$, and it suffices to show that $\mathrm{parent}_{N'}(x_i) = u'_i$ for each $1 \leq i \leq k-1$, as $\mathrm{parent}_{N'}(x_k) = \mathrm{parent}_{N'}(v) = u'_k$ is known. Additionally, we will show that $t(u_i) = t'(u'_i)$ for all $1 \leq i < k$. We proceed by induction.

First consider the vertex x_1. By Lemma 2.1 we have $\mathrm{lca}_N(x_1x_i) = u_1$ for all $1 < i \leq k$. Hence

$$\sigma(x_1x_i) = t(u_1) =: \bullet \tag{11}$$

follows for all $1 < i \leq k$. For contradiction, assume $\mathrm{parent}_{N'}(x_1) = u'_i$ for $i \neq 1$. Since Σ is a polar-cat we know $|V(\Omega_1)| = k \geq 2$. Hence there is at least one leaf, namely $x_k = v$ such that $u'_i \succ_{N'} \mathrm{parent}_{N'}(x_k)$. By (11) and Lemma 2.1 it thus follows that $t'(u'_i) = t'(\mathrm{lca}_{N'}(x_1x_k)) = \bullet$. Since $i \neq 1$ and N' is quasi-discriminating there must exist some u'_j such that $u'_j \succ_{N'} u'_i$ and such that $t'(u'_j) \neq \bullet$. Let x_m be the leaf-child of u'_j. By Lemma 2.1 we thus have $\sigma(x_1x_m) = t'(\mathrm{lca}_{N'}(x_1x_m)) = t'(u'_j) \neq \bullet$, which contradicts (11). We conclude that $\mathrm{parent}_{N'}(x_1) = u'_1$ and that $t'(u'_1) = t(u_1)$.

Now assume that there is some index $2 \leq j < k$ such that $\mathrm{parent}_{N'}(x_i) = u'_i$ and $t(u_i) = t'(u'_i)$ for all $1 \leq i < j$. Consider the vertex x_j. By Lemma 2.1

we have $\mathrm{lca}_N(x_j x_i) = u_j$ for all x_i such that $j < i \leq k$. Hence

$$\sigma(x_j x_i) = t(u_j) =: \blacktriangle \tag{12}$$

follows for all $j < i \leq k$. For contradiction, assume $\mathrm{parent}_{N'}(x_j) = u'_m \neq u'_j$. By assumption, $m > j$. However $m < k$, since $x_j \neq v$, and the only leaf-child of u'_k is v. Moreover, $u'_m \succ_{N'} \mathrm{parent}_{N'}(x_k)$. By (12) and Lemma 2.1 we thus have $t'(u'_m) = t'(\mathrm{lca}_{N'}(x_j x_k)) = \blacktriangle$. Now consider the vertex u'_{m-1}, i.e. the parent of u'_m (in N'). Since N' is quasi-discriminating, we must have that $t'(u'_{m-1}) \neq \blacktriangle$. Denote the unique leaf-child of u'_{m-1} by x_l. By the induction hypothesis, we must have that $l > j$. However, Lemma 2.1 and $u'_{m-1} \succ_{N'} u'_m$ implies that $\mathrm{lca}_{N'}(x_j x_l) = \mathrm{parent}_{N'}(x_l) = u'_{m-1}$, so that $\sigma(x_j x_l) = t'(u'_{m-1}) \neq \blacktriangle$. This contradicts (12), so that we may conclude that $\mathrm{parent}_{N'}(x_j) = u'_j$. Additionally, we already established that $t'(u'_j) = t(u_j)$. By the principle of induction we thus have that $\mathrm{parent}_{N'}(x_i) = u'_i$ and $t(u_i) = t'(u'_i)$ for all $1 \leq i < k$.

The same argument may be applied to the other side of N and N'. More precisely, if $Q = w_1 w_2 \ldots w_l \neq P$ is the other side of the underlying cycle of N and $Q' = w'_1 w'_2 \ldots w'_l \neq P'$ is the other side of the underlying cycle of N', then the vertices of Ω_2 may be ordered as y_1, y_2, ..., y_l so that $\mathrm{parent}_N(y_i) = w_i$ and $\mathrm{parent}_{N'}(y_i) = w'_i$ for all $1 \leq i \leq l$. Additionally, $t(w_i) = t'(w'_i)$ holds for all $1 \leq i < l$. We may thus define a bijective map $\varphi : V(N) \to V(N')$ by

$$\varphi(x) := \begin{cases} x & \text{if } x \in \mathcal{L}(N) \\ u'_i & \text{if } x = u_i \text{ for some } 1 \leq i \leq k \\ w'_i & \text{if } x = w_i \text{ for some } 1 \leq i \leq l \\ \rho_{N'} & \text{if } v = \rho_N. \end{cases}$$

Clearly, φ is an isomorphism of N and N' as directed graphs. In particular $\mathcal{L}(N) = \mathcal{L}(N')$ and $\varphi(x) = x$ for all $x \in \mathcal{L}(N)$. Additionally, since $t(u_i) = t'(u'_i)$ for all $1 \leq i < k$, $t(w_i) = t'(w'_i)$ for all $1 \leq i < l$ and the roots of N and N' must be labeled with the join-color of Σ by t and t' respectively, we have that $t(u) = t'(\varphi(u))$ for all vertices u except, possibly, $u = \eta_N$. By Observation 2.1 we thus have $t(u) = t(u')$ if and only if $t'(\varphi(u)) = t'(\varphi(u'))$ for all $u, u' \in V^0_{\mathrm{lca}}(N)$. In other words, φ is an isomorphism of (N, t) and (N', t'), and there is thus a unique galled tree that explains Σ. As t and t' agree on all inner vertices except possibly η_N, the labeling t is unique up to the label of the hybrid vertex of N. $\square$

Although Theorem 4.6 gives us a valuable connection between having a fixpoint and uniqueness of the galled tree that explains a polar-cat, the property of having a fixpoint is yet not a very helpful concept. The rest of this section is thus concerned with finding sufficient conditions for when a polar-cat has a fixpoint.

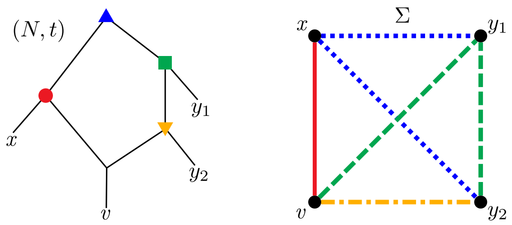


Figure 28: The depicted $(v, \Sigma[\{v,x\}], \Sigma[\{v,y_1,y_2\}])$-polar-cat Σ is no (w, Ω_1, Ω_2)-polar-cat for $w \neq v$, although it satisfies neither condition (C1) nor (C2). In other words, it has a fixpoint.

For the case of two colors, the property of a polar-cat having a fixpoint was characterized with a condition called being well-behaved which, very roughly, ensures the polar-cat in question is "large enough" [19, Def. 6.11, Lem. 6.12]. Formally, an edge-colored (v, Ω_1, Ω_2)-polar-cat with two edge-colors is *well-behaved* if one of the following conditions is satisfied

(C1) $|V(\Omega_1)| \geq 3$ and $|V(\Omega_2)| \geq 3$, or

(C2) $|V(\Omega_i)| = 2$ and $|V(\Omega_j)| \geq 5$ for distinct $i, j \in \{1, 2\}$.

We will soon see that for more than two edge-colors, these two conditions are indeed sufficient for having a fixpoint, but it is no longer a necessity.

To see the latter, consider the edge-colored graph Σ depicted in Figure 28. One easily verifies that Σ is a (v, Ω_1, Ω_2)-polar-cat for $\Omega_1 = \Sigma[\{v, x\}]$ and $\Omega_2 = \Sigma[\{v, y_1, y_2\}]$, for example by verifying that the elementary, quasi-discriminating and strong galled tree (N, t) in Figure 28 explains Σ. It cannot be a (y_i, Π_1, Π_2)-polar-cat for $i = 1$ or $i = 2$ since neither $\Sigma - y_1$ nor $\Sigma - y_2$ is $\mathfrak{unp}$ (cf. Prop. 3.8). Moreover, it is not a (x, Π_1, Π_2)-polar-cat since

$$\Sigma - x = \Sigma[\{y_1\}] \otimes_{\blacksquare} \Sigma[\{v, y_2\}]$$

would enforce $\Pi_i = \Sigma[\{x, v, y_2\}]$ for either $i = 1$ or $i = 2$, and $\Sigma[\{x, v, y_2\}]$ is certainly not caterpillar-explainable (it is a rainbow triangle!). That is, Σ has a fixpoint, although neither condition (C1) nor (C2) is satisfied.

Proposition 4.7. *If a polar-cat Σ satisfies condition (C1) or (C2), then it has a fixpoint. In particular, there is a unique galled tree that explains Σ.*

Proof. Let $\Sigma = (V, \sigma)$ be a (v, Ω_1, Ω_2)-polar-cat satisfying condition (C1) or (C2). Since the case of two colors was proved in [19, Lem. 6.12], we may assume Σ has at least three edge-colors. By Proposition 4.4 there are vertices $x, y \in V$ such that $\Sigma[\{x, y, v\}]$ is a rainbow triangle, where $x \in V(\Omega_1)$ and $y \in V(\Omega_2)$. Let $\bullet = \sigma(xy)$ and $\blacktriangle = \sigma(vy)$ be two of the (necessarily distinct) edge-colors of $\Sigma[\{x, y, v\}]$.

For contradiction, assume Σ additionally is a (w, Π_1, Π_2)-polar-cat for some $w \neq v$. With Proposition 4.4 in mind, w must be a vertex of the rainbow triangle $\Sigma[\{x, y, v\}]$, while the two vertices in $\{x, y, v\} \setminus \{w\}$ belong to Π_1 and Π_2, respectively. We thus assume, without loss of generality, that $w = x$ and that $v \in V(\Pi_1)$ while $y \in V(\Pi_2)$.

By definition of polar-cats, it is easily verified that

$$\Sigma - v = (\Omega_1 - v) \otimes_{\color{red}\bullet} (\Omega_2 - v)\,, \text{ and} \tag{13}$$
$$\Sigma - x = (\Pi_1 - x) \otimes_{\color{blue}\blacktriangle} (\Pi_2 - x)\,, \tag{14}$$

for the (earlier defined) colors ${\color{red}\bullet}$ and ${\color{blue}\blacktriangle}$. The most important direct consequence of (13) and (14) are the following two implications, which we will make use of multiple times. For any two vertices $a, b \in V \setminus \{v\}$ we have that

$$\sigma(a, b) \neq {\color{red}\bullet} \implies a, b \in V(\Omega_i) \text{ for } i = 1 \text{ or } i = 2, \tag{15}$$

while for any two vertices $a, b \in V \setminus \{x\}$ we have that

$$\sigma(a, b) \neq {\color{blue}\blacktriangle} \implies a, b \in V(\Pi_i) \text{ for } i = 1 \text{ or } i = 2. \tag{16}$$

We first assume that condition (C1) is satisfied, that is, assume that $|V(\Omega_1)| \geq 3$ and $|V(\Omega_2)| \geq 3$. Hence there exists some $x' \in V(\Omega_1) \setminus \{x, v\}$ and some $y' \in V(\Omega_2) \setminus \{y, v\}$. In fact, since Ω_2 is explained by a discriminating caterpillar we may choose y' such that $\sigma(vy') \neq \sigma(vy)$, that is, such that $\sigma(vy') \neq {\color{blue}\blacktriangle}$. Thus (16) and $v \in V(\Pi_1)$ implies that $y' \in V(\Pi_1)$. Now, since $x' \in V(\Omega_1) \setminus \{v\}$ and $y' \in V(\Omega_2) \setminus \{v\}$ we have, by (13), that $\sigma(x'y') = {\color{red}\bullet} \neq {\color{blue}\blacktriangle}$. Hence $y' \in V(\Pi_1)$ and (16) implies $x' \in V(\Pi_1)$. To summarize, we have shown that $x' \in V(\Pi_1) \setminus \{x\}$, and know by assumption respectively (13) that $y \in V(\Pi_2) \setminus \{x\}$ and $\sigma(x'y) = {\color{red}\bullet}$. This is a contradiction to (16). We thus conclude that if Σ satisfies condition (C1) the it has a fixpoint.

Condition (C2) can be satisfied in two ways, depending on wether it is Ω_1 or Ω_2 that contains two vertices. We need to consider both cases. First assume that $|V(\Omega_1)| = 2$ i.e. $V(\Omega_1) = \{x, v\}$ and that $|V(\Omega_2)| \geq 5$. Hence there are $k \geq 3$ vertices z_1, z_2, ..., z_k such that $V(\Omega_2) = \{y, z_1, \ldots, z_k, v\}$. By (13), we must have

$$\sigma(xz_i) = {\color{red}\bullet} \text{ for } i = 1, \ldots, k. \tag{17}$$

Let $I \subseteq \{1, \ldots, k\}$ be the subset of indices such that $z_i \in V(\Pi_2)$ if and only if $i \in I$. Since $x, y \in V(\Pi_2)$ but $v \notin V(\Pi_2)$, we thus have that $|V(\Pi_2)| = |I| + 2$. Note that by definition of polar-cats, x is part of the cherry of the discriminating caterpillar tree that explains Π_1, so we may apply Lemma 4.3. By (17), $\sigma(xz_i) = \sigma(xz_j) = \sigma(xy)$ for all $i, j \in I$, from which the lemma enforces that $|I| + 1 \leq \lfloor (|I| + 2)/2 \rfloor$. As the reader may verify easily, this is true only if $|I| = 0$, that is, only if $V(\Pi_1) = \{x, v, z_1, \ldots, z_k\}$, so that $|V(\Pi_1)| = k + 2$. Since x is also part of the cherry of the discriminating caterpillar tree that explains Π_1 and, again by (17), we have $\sigma(xz_i) = \sigma(xz_j)$ for all k indices $i, j \in \{1, \ldots, k\}$ and Lemma 4.3 thus implies that $k \leq \lfloor (k + 2)/2 \rfloor$. This inequality may again be verified with standardized methods to imply $k \leq 2$, which contradicts $|V(\Omega_2)| \geq 5$.

Lastly, we instead assume that $V(\Omega_1) = \{x, z_1, \ldots, z_k, v\}$ for $k \geq 3$ vertices z_1,...,z_k, so that $|V(\Omega_1)| \geq 5$. Additionally assume $V(\Omega_2) = \{v, y\}$ so that condition (C2) is satisfied. On one hand, (13) implies that $\sigma(yz_i) = \bullet$ for $i = 1, 2, \ldots, k$ from which (16) and $y \in V(\Pi_2)$ implies $z_i \in V(\Pi_2) \setminus \{x\}$ for $i = 1, 2, \ldots, k$. The latter taken together with $v \in V(\Pi_1) \setminus \{x\}$ and (14) implies $\sigma(vz_i) = \blacktriangle$ for $i = 1, 2, \ldots, k$. Now note that v is part of the cherry of the discriminating caterpillar tree that explains Ω_1, so we may apply Lemma 4.3 once again. There are k edges $\{v, z_i\}$ in the same color and $|V(\Omega_1)| = k + 2$, hence $k \leq \lfloor (k + 2)/2 \rfloor$ is implied by the lemma. Once again we get the contradiction of $k \leq 2$. Hence we conclude that if Σ satisfies condition (C2) then it has a fixpoint. By Theorem 4.6, there is thus a unique galled tree that explains Σ. □

It turns out that another sufficient condition for a polar-cat having a fixpoint instead depend on the number of edge-colors, rather than the number of vertices in the two subgraphs. More precisely, we have the following proposition.

Proposition 4.8. *If a polar-cat Σ has at least four edge-colors, then it has a fixpoint. In particular, there is a unique galled tree that explains Σ.*

Proof. We begin by repeating the first argument of the proof of Proposition 4.7. If $\Sigma = (V, \sigma)$ is a (v, Ω_1, Ω_2)-polar-cat with at least four edge-colors, then there exists (by Proposition 4.4) vertices x and y such that $\Sigma[\{v, x, y\}]$ is a rainbow triangle. Additionally, $x \in V(\Omega_1)$ and $y \in V(\Omega_2)$. We may introduce the (necessarily distinct) colors $\bullet := \sigma(xy)$, $\blacktriangle := \sigma(vy)$ and $\blacksquare := \sigma(xv)$.

Since there are at least four edge-colors in Σ, there are at least four vertices in Σ. Note that v is part of the cherry of the respective caterpillar tree that explains Ω_1 and Ω_2, so by Lemma 4.3 there in particular exist some $z \notin \{v, x, y\}$ such that $\sigma(vz) = \blacktriangledown \notin \{\bullet, \blacktriangle, \blacksquare\}$. We may, without loss of generality, assume that $z \in V(\Omega_2)$. Since $\Sigma - v = (\Omega_1 - v) \otimes_{\bullet} (\Omega_2 - v)$ follows from definition, and $x \in V(\Omega_1)$ it thus follows that $\sigma(xz) = \bullet$. We have thus established the edge-color of all edges except $\{y, z\}$ of the subgraph $\Sigma[\{v, x, y, z\}]$: see Figure 29 for a visual reminder.

First note that $\Sigma[\{v, x, z\}]$ is a rainbow triangle and a subgraph of $\Sigma - y$. Similarly, $\Sigma[\{v, x, y\}]$ is a rainbow triangle and a subgraph of $\Sigma - u$ for any $u \in V \setminus \{v, x, y\}$. By Observation 2.2 rainbow triangles are primitive, so that Theorem 2.5 implies that neither $\Sigma - y$ nor $\Sigma - u$ (for any $u \in V \setminus \{v, x, y\}$) is

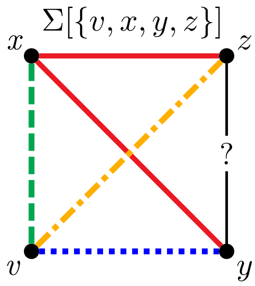


Figure 29: The subgraph $\Sigma[\{v, x, y, z\}]$ always appear in a (v, Ω_1, Ω_2)-polar-cat with at least four edge-colors, where $\bullet$, $\blacktriangle$, $\blacksquare$ and $\blacktriangledown$ are distinct colors. Only the edge-color of the edge $\{y, z\}$ may vary. For further details, see proof of Proposition 4.8.

$\mathfrak{unp}$. Hence Proposition 3.8 ensures that Σ cannot be a (u, Π_1, Π_2)-polar-cat for $u = y$ or for any $u \in V \setminus \{v, x, y\}$, however the subgraphs Π_1 and Π_2 are chosen.

It remains to show that Σ cannot be a (x, Π_1, Π_2)-polar-cat. The argument depends on the color of $\{y, z\}$, and we have three possible cases: $\sigma(yz) = \blacktriangle$, $\sigma(yz) = \blacktriangledown$ or $\sigma(yz) \notin \{\blacktriangle, \blacktriangledown\}$. If the latter holds, then $\Sigma[\{v, y, z\}]$ is a rainbow triangle and a subgraph of $\Sigma - x$, implying that $\Sigma - x$ is not $\mathfrak{unp}$. Hence, by Proposition 3.8, Σ is not a (x, Π_1, Π_2)-polar-cat.

Secondly, assume $\sigma(yz) = \blacktriangle$. If Σ where a (x, Π_1, Π_2)-polar-cat for some subgraphs Π_1 and Π_2, then there must be some color k such that

$$\Sigma - x = (\Pi_1 - x) \otimes_k (\Pi_2 - x). \tag{18}$$

In particular $\Sigma - x$ contains the subgraph $\Sigma[\{v, y, z\}]$, for which $\Sigma[\{v, y, z\}] = \Sigma[\{v, z\}] \otimes_{\blacktriangle} \Sigma[\{y\}]$ holds. For this to be true at the same time as (18), we must have that $v, z \in V(\Pi_i)$ for either $i = 1$ or $i = 2$, while $y \in V(\Pi_j)$ for $j \in \{1, 2\} \setminus \{i\}$. Since $x \in V(\Pi_i)$ as well, Π_i contains the subgraph $\Sigma[\{x, v, z\}]$, which is a rainbow triangle. Thus Π_i cannot possibly be explained by a caterpillar tree, and Σ is not a (x, Π_1, Π_2)-polar-cat.

The case when $\sigma(yz) = \blacktriangledown$ is almost identical to when $\sigma(yz) = \blacktriangle$ — the subgraph Π_i will contain the rainbow triangle $\Sigma[\{x, v, y\}]$ for either $i = 1$ or $i = 2$ — so we omit the details. We have thus shown that if a polar-cat has at least four edge-colors, then it has a fixpoint. By Theorem 4.6, there is thus a unique galled tree that explains Σ. □

We conclude this section by stating the following result for later reference, which is an immediate consequence of Theorem 4.6, Proposition 4.7 and Proposition 4.8.

Proposition 4.9. *Let Σ be a polar-cat. If Σ has at least four edge-colors or satisfies condition (C1) or (C2), then there is a unique galled tree (N, t) that explains Σ. Additionally, the labeling map t of (N, t) is uniquely determined, up to the label of the hybrid vertex of N.*

We have thus given three sufficient conditions for a polar-cat being explained by a unique galled tree. We now consider the reverse question: when are there at least two distinct galled trees that explain a polar-cat?

4.3 Non-uniqueness

Due to Proposition 4.9 a polar-cat Σ may only be explained by two or more distinct galled trees if it has at most three edge-colors and fails both condition (C1) and condition (C2). For ease of discussion, we note that a (v, Ω_1, Ω_2)-polar-cat fails all these three conditions if and only if it

(C3) has two or three edge-colors and $|V(\Omega_i)| = 2$ while $2 \leq |V(\Omega_j)| \leq 4$ for distinct $i, j \in \{1, 2\}$,

as $|V(\Omega_1)| \geq 2$ and $|V(\Omega_2)| \geq 2$ are ensured by definition (and there are no polar-cats with only one edge-color, as polar-cats are primitive). Still, as we will show in this section, there are polar-cats satisfying condition (C3) that are still explained by a unique galled tree. That is, the converse of Proposition 4.9 is not true. We have not been able to formulate a succinct condition that ensures a polar-cat may be explained by multiple galled trees. Instead we revert to a full enumeration of all polar-cats that that satisfies condition (C3), possible since condition (C3) in particular implies that the polar-cat in question has at most five vertices.

We first limit the scope of our search in two ways. As we have noted earlier (Observation 3.1) the only polar-cat on three vertices is a rainbow triangle. In Example 4.1 we gave two distinct galled trees that explained (the same) rainbow triangle. We may thus consider the case of three vertices to be resolved: there is only one such polar-cat, and it is not explained by a unique galled tree.

Secondly we note that the converse of Proposition 4.9 was proved for the case of polar-cats with two edge-colors in [19, Prop. 6.13]. That is, if a polar-cat Σ satisfies condition (C3) and has precisely two edge-colors, then there are necessarily distinct galled trees that explain Σ. In fact, there's only two possible such polar-cats: one with four vertices, and one with five. For completeness' sake we depict these two edge-colored graphs in Figure 30, together with distinct galled trees that explain them. See [19] for further details.

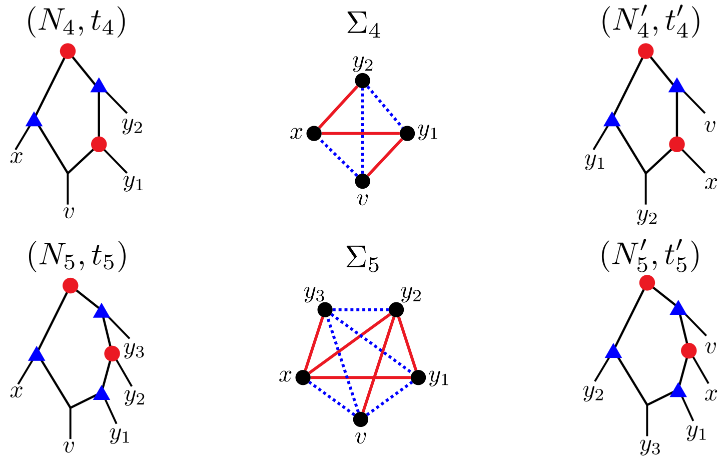


Figure 30: The two polar cats Σ_4 and Σ_5 on four and five vertices respectively, that satisfies condition (C3) and have only two edge-colors. Additionally we give four distinct galled trees (N_4, t_4), (N'_4, t'_4), (N_5, t_5) and (N'_5, t'_5). The reader may verify that $\mathcal{G}(N_4, t_4) = \Sigma_4 = \mathcal{G}(N'_4, t'_4)$ and $\mathcal{G}(N_5, t_5) = \Sigma_5 = \mathcal{G}(N'_5, t'_5)$.

We can thus restrict ourself further, by only considering (v, Ω_1, Ω_2)-polar-cats that

(C3') has three edge-colors and $|V(\Omega_i)| = 2$ while $|V(\Omega_j)| \in \{3, 4\}$ for distinct $i, j \in \{1, 2\}$.

By Theorem 3.2 there exist an elementary, quasi-discriminating and strong galled tree (N, t) that explains a given (v, Ω_1, Ω_2)-polar-cat Σ satisfying condition (C3'). Without loss of generality, assume $V(\Omega_1) = \{v, x\}$ while $V(\Omega_2) = \{v, y_1, \ldots, y_k\}$, where $k = 2$ or $k = 3$. Additionally, Lemma 3.3 ensures that the vertices of Ω_1 belong to one of the sides of the underlying cycle of N, while the vertices of Ω_2 belong to the other side. That is, by condition (C3'), v and x belong to one of the sides of the underlying cycle of N, while v, y_1, ..., and y_k belong the other side. Per construction, v is the child of the unique hybrid η of N. Again without loss of generality, we may further assume that $\text{parent}(y_2) \succ_N \text{parent}(y_1)$ and, in case $k = 3$, that $\text{parent}(y_3) \succ_N \text{parent}(y_2)$. We have thus constructed two galled trees $N_{(4)}$ and $N_{(5)}$ depicted in Figure 31 and proved the following.

Observation 4.1. If Σ is a polar-cat that satisfies condition $(C3')$ and thus, has $n \in \{4, 5\}$ vertices, then there is a quasi-discriminating labeling t of $N_{(n)}$ such that $\mathcal{G}(N_{(n)}, t) \simeq \Sigma$.

In fact, even the following stronger statement is true.

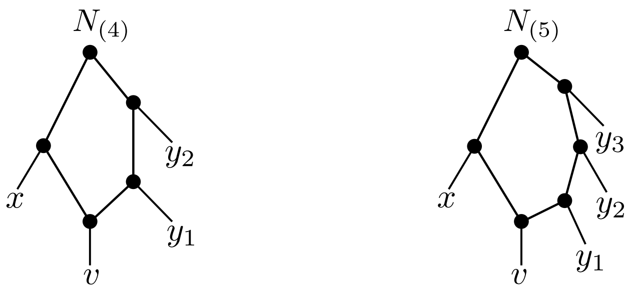


Figure 31: The two galled trees $N_{(4)}$ and $N_{(5)}$.

Proposition 4.10. *If Σ is a polar-cat that satisfies condition (C3') and thus, has $n \in \{4,5\}$ vertices, then there is a quasi-discriminating labeling t of $N := N_{(n)}$ such that $t(\text{parent}(x)) = \blacktriangle$, $t(\rho_N) = \bullet$, $t(u) = \blacksquare$ for some vertex u of N, and $\mathcal{G}(N_{(n)}, t) \simeq \Sigma$.*

Proof. Let Σ be a polar-cat that satisfies condition (C3'). Suppose Σ has n vertices and put $N := N_{(n)}$. By Observation 4.1 there is a labeling t of N such that $\mathcal{G}(N,t) \simeq \Sigma$. We construct another labeling t' of N by putting

$$t'(u) := \begin{cases} \blacktriangle & \text{if } t(u) = t(\text{parent}(x)) \\ \bullet & \text{if } t(u) = t(\rho_N) \\ \blacksquare & \text{if } t(u) \notin \{t(\text{parent}(x)), t(\rho_N)\} \text{ and } u \notin \mathcal{L}(N) \\ \odot & \text{if } u \in \mathcal{L}(N) \end{cases}$$

for each vertex u of N. Note that by condition (C3') respectively construction it follows that $\mathcal{G}(N,t)$ and $\mathcal{G}(N,t')$ (respectively) has precisely three edge-colors. Thus the definition of t' ensures that for vertices $u, u' \in V(N)$ we have

$$t(u) = t(u') \text{ if and only if } t'(u) = t'(u'). \tag{19}$$

The definition of t' ensures that $t'(\text{parent}(x)) = \blacktriangle$ and $t'(\rho_N) = \bullet$. Furthermore, since $\mathcal{G}(N,t)$ has three edge-colors, there must be some internal vertex u of N with $t(u) \notin \{t(\text{parent}(x)), t(\rho_N)\}$, hence there exists some vertex u such that $t'(u) = \blacksquare$. It thus remains only to show that $\mathcal{G}(N,t) \simeq \mathcal{G}(N,t')$.

For brevity, introduce $V := \mathcal{L}(N)$ and the edge-coloring maps σ and σ' of $\mathcal{G}(N,t)$ and $\mathcal{G}(N,t')$, respectively, so that $\mathcal{G}(N,t) = (V,\sigma)$ and $\mathcal{G}(N,t') = (V,\sigma')$. To conclude that $\mathcal{G}(N,t) \simeq \mathcal{G}(N,t')$ we show that

$$\sigma(uw) = \sigma(u'w') \iff \sigma'(uw) = \sigma'(u'w')$$

for all $u, w, u', w' \in V$ such that $u \neq w$ and $u' \neq w'$, i.e. we show that the identity map on $\binom{V}{2}$ is an isomorphism of $\mathcal{G}(N,t)$ and $\mathcal{G}(N,t')$.

Note that for any distinct $u, w \in V$ we have $t(\text{lca}(u,w)) = \sigma(uw)$ and $t'(\text{lca}(u,w)) = \sigma'(uw)$. That is, $\sigma(uw) = \sigma(u'w')$ is satisfied if and only if $t(\text{lca}(u,w)) = t(\text{lca}(u',w'))$. By (19), $t(\text{lca}(u,w)) = t(\text{lca}(u',w'))$ holds if and only if $t'(\text{lca}(u,w)) = t'(\text{lca}(u',w'))$, which is equivalent to $\sigma'(uw) = \sigma(u'w')$. We thus conclude that $\mathcal{G}(N,t) \simeq \mathcal{G}(N,t')$, from which the statement follows. $\square$

As a result of Proposition 4.10 we may first enumerate all possible quasi-discriminating labelings t_i of $N_{(n)}$ for $n = 4$ and $n = 5$ where $t_i(\text{parent}(x)) = \blacktriangle$ and $t_i(\rho_{N_{(n)}}) = \bullet$. For ease of discussion, we call such a labeling of $N_{(n)}$ a $(\blacktriangle, \bullet)$-*fixed labeling*. We then consider the edge-colored graphs $\mathcal{G}(N_{(n)}, t_i)$. The resulting polar-cats with three edge-colors are the only polar-cats (up to isomorphism) that satisfy condition (C3'). To generate quasi-discriminating labelings of $N_{(4)}$ and $N_{(5)}$ we note that the labels of $\text{parent}(y_1)$, $\text{parent}(y_2)$ and, possibly, $\text{parent}(y_3)$ are the only three labels that are to be determined for each $(\blacktriangle, \bullet)$-fixed labeling t_i. To this end, consider an ordered tuple $(c_1, c_2, \ldots, c_k)$ of colors where $k \in \{4,5\}$, $c_1 = \blacktriangle$, $c_2 = \bullet$ and $c_i \neq c_{i+1}$ for $i = 1, \ldots, k-1$. We call such a tuple a *labeling array* or, if k needs to be emphasized, a k-*labeling array*. Clearly, a $(\blacktriangle, \bullet)$-fixed quasi-discriminating labeling t of $N_{(k)}$, for $k \in \{4,5\}$, can be constructed from a k-labeling array $(c_1, c_2, \ldots, c_k)$ by defining

$$t(u) := \begin{cases} \blacktriangle & \text{if } u = \text{parent}(x) \\ \bullet & \text{if } u = \rho_{N_{(k)}} \\ c_i & \text{if } u = \text{parent}(y_i) \\ \odot & \text{if } u \text{ is a leaf} \end{cases}$$

for each vertex u of $N_{(k)}$. For the hybrid vertex η of $N_{(k)}$ the label may be arbitrarily defined. Conversely, given a $(\blacktriangle, \bullet)$-fixed quasi-discriminating labeling t of $N_{(n)}$, there is a corresponding n-labeling array $(c_1, \ldots, c_n)$ defined by putting $c_1 = \blacktriangle$, $c_2 = \bullet$ and $c_i := t(\text{parent}(y_i))$ for $i = 1, \ldots, n-2$. Intuitively speaking, a labeling array encodes the labels of the vertices of the underlying cycle of $N_{(n)}$ starting at $\text{parent}(x)$ and proceeding clock-wise. Stated more succinctly, we have the following.

Observation 4.2. There is a 1-to-1 correspondence between 4-labeling arrays and $(\blacktriangle, \bullet)$-fixed quasi-discriminating labelings s_i of $N_{(4)}$. Similarly, there is a 1-to-1 correspondence between 5-labeling arrays and $(\blacktriangle, \bullet)$-fixed quasi-discriminating labelings t_i of $N_{(5)}$.

To give an example, the 5-labeling array $(\blacktriangle, \bullet, \blacktriangle, \bullet, \blacktriangle)$ corresponds to the labeling t_5 of $N_{(5)}$ depicted in Figure 30 (note $N_5 = N_{(5)}$).

The algorithm `genLA` provides a way of generating distinct k-labeling arrays; we are interested only in the case of $k = 4$ and $k = 5$. The procedure is rather simple: at each step we extend a set of i-labeling arrays

Algorithm 1: The algorithm `genLA`, used for generating labeling arrays.

```
  Input: Integer k ≥ 3, array N of new colors.
  Output: A set of all possible labeling arrays of length k.
1 begin
2   currentLAs ← array with the single labeling array (▲, ●)
3   for 1, 2, ..., k − 2 do
4     longerLAs ← empty array
5     for LA in currentLAs do
6       Extend LA with next color of N not in LA and add to
          longerLAs
7       for each used color c except the last element of LA do
8         Extend LA with c and add to longerLAs
9     Replace currentLAs with longerLAs
10  return currentLAs
```

to a larger set of $(i+1)$-labeling arrays by, for each i-labeling array L, (1) extending L by introducing a color not yet appearing in L, and (2) extending L by repeating any one color appearing in L, as long as it is not same color as the last color of L. Step (1) is performed on row 7, while step (2) is performed on rows 8–9. It is clear that after each iteration i, the array `currentLAs` will contain only $(i+2)$-labeling arrays, thus the output will be an array containing k-labeling arrays, where k is given as output. As a smaller technicality, we assume that new colors are introduced in the order ■, ▼ and ⬢ i.e. that the `genLA` is called with the array N = [■, ▼, ⬢]. This is required to align our output with Proposition 4.10. We visualize the procedure for 4- and 5-labeling arrays in Figure 32 and Figure 33, respectively. Moreover, we provide an implementation of `genLA` (in Python) hosted at (`https://github.com/AnnaLindeberg/kLabelingArrays` [25]), although the two figures should suffice for most readers. We collect the above discussion in a lemma, for later reference.

Lemma 4.11. `genLA`(k, [■, ▼, ⬢]) *correctly outputs the set of all k-labeling arrays, where $k \in \{4, 5\}$. In particular, the output will include every k-labeling array whose elements are all members of the set* {▲, ●, ■}.

A call to `genLA`(4, [■, ▼, ⬢]) yields five 4-labeling arrays, namely:

$L_1^{(4)}$ = (▲, ●, ▲, ●), $L_2^{(4)}$ = (▲, ●, ▲, ■), $L_3^{(4)}$ = (▲, ●, ■, ▲),
$L_4^{(4)}$ = (▲, ●, ■, ●), and $L_5^{(4)}$ = (▲, ●, ■, ▼),

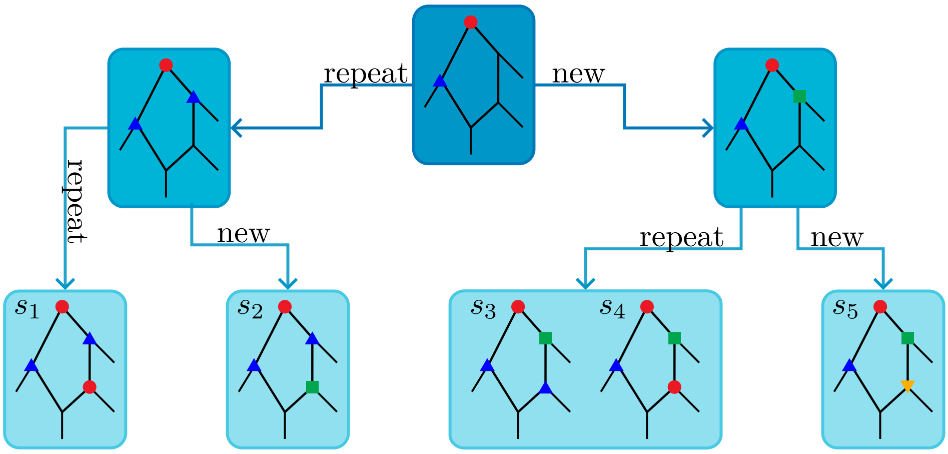


Figure 32: Enumeration of all (▲, ●)-fixed quasi-discriminating labelings s_i of $N_{(4)}$. The lighter the bounding box, the fewer vertices are yet to be labeled. In the end, there are five distinct s_i, out of which three have precisely three distinct labels. For readability, the leaves are not labeled but the underlying tree is still $N_{(4)}$ of Figure 31.

while a call to `genLA`(5, [■, ▼, ⬢]) yields the following fifteen 5-labeling arrays:

$$
\begin{aligned}
L_1^{(5)} &= (\blacktriangle, \bullet, \blacktriangle, \bullet, \blacktriangle),\ L_2^{(5)} = (\blacktriangle, \bullet, \blacktriangle, \bullet, \blacksquare),\ L_3^{(5)} = (\blacktriangle, \bullet, \blacktriangle, \blacksquare, \bullet),\\
L_4^{(5)} &= (\blacktriangle, \bullet, \blacktriangle, \blacksquare, \blacktriangle),\ L_5^{(5)} = (\blacktriangle, \bullet, \blacktriangle, \blacksquare, \blacktriangledown),\ L_6^{(5)} = (\blacktriangle, \bullet, \blacksquare, \blacktriangle, \bullet),\\
L_7^{(5)} &= (\blacktriangle, \bullet, \blacksquare, \blacktriangle, \blacksquare),\ L_8^{(5)} = (\blacktriangle, \bullet, \blacksquare, \blacktriangle, \blacktriangledown),\ L_9^{(5)} = (\blacktriangle, \bullet, \blacksquare, \bullet, \blacktriangle),\\
L_{10}^{(5)} &= (\blacktriangle, \bullet, \blacksquare, \bullet, \blacksquare),\ L_{11}^{(5)} = (\blacktriangle, \bullet, \blacksquare, \bullet, \blacktriangledown),\ L_{12}^{(5)} = (\blacktriangle, \bullet, \blacksquare, \blacktriangledown, \blacktriangle),\\
L_{13}^{(5)} &= (\blacktriangle, \bullet, \blacksquare, \blacktriangledown, \bullet),\ L_{14}^{(5)} = (\blacktriangle, \bullet, \blacksquare, \blacktriangledown, \blacksquare), \text{ and } L_{15}^{(5)} = (\blacktriangle, \bullet, \blacksquare, \blacktriangledown, \text{⬢})
\end{aligned}
$$

Note that Figure 32 and Figure 33 presents these five respectively fifteen labeling arrays by their counterpart as a (▲, ●)-fixed quasi-discriminating labeling of $N_{(n)}$ — as allowed by Observation 4.2. Formally speaking, we put s_i as the (▲, ●)-fixed quasi-discriminating labeling of $N_{(4)}$ that corresponds to $L_i^{(4)}$ for each $i = 1, \dots, 5$. Similarly, t_i is the (▲, ●)-fixed quasi-discriminating labeling of $N_{(5)}$ that corresponds to $L_i^{(5)}$ for each $i = 1, \dots, 15$. See also Table 1 and Table 2.

Now considering the polar-cats $\mathcal{G}(N_{(4)}, s_i)$ respectively $\mathcal{G}(N_{(5)}, t_i)$ and taking the number of edge-colors into account there are, as easily verified by the reader, three respectively seven polar-cats that satisfy condition (C3'). On four vertices, these are $\mathcal{G}(N_{(4)}, s_2)$, $\mathcal{G}(N_{(4)}, s_3)$ and $\mathcal{G}(N_{(4)}, s_4)$. Those with five vertices are $\mathcal{G}(N_{(5)}, t_2)$, $\mathcal{G}(N_{(5)}, t_3)$, $\mathcal{G}(N_{(5)}, t_4)$, $\mathcal{G}(N_{(5)}, t_6)$, $\mathcal{G}(N_{(5)}, t_7)$, $\mathcal{G}(N_{(5)}, t_9)$ and $\mathcal{G}(N_{(5)}, t_{10})$. All other s_i and t_i yield edge-colored graphs either with two edge-colors only, or with more than three

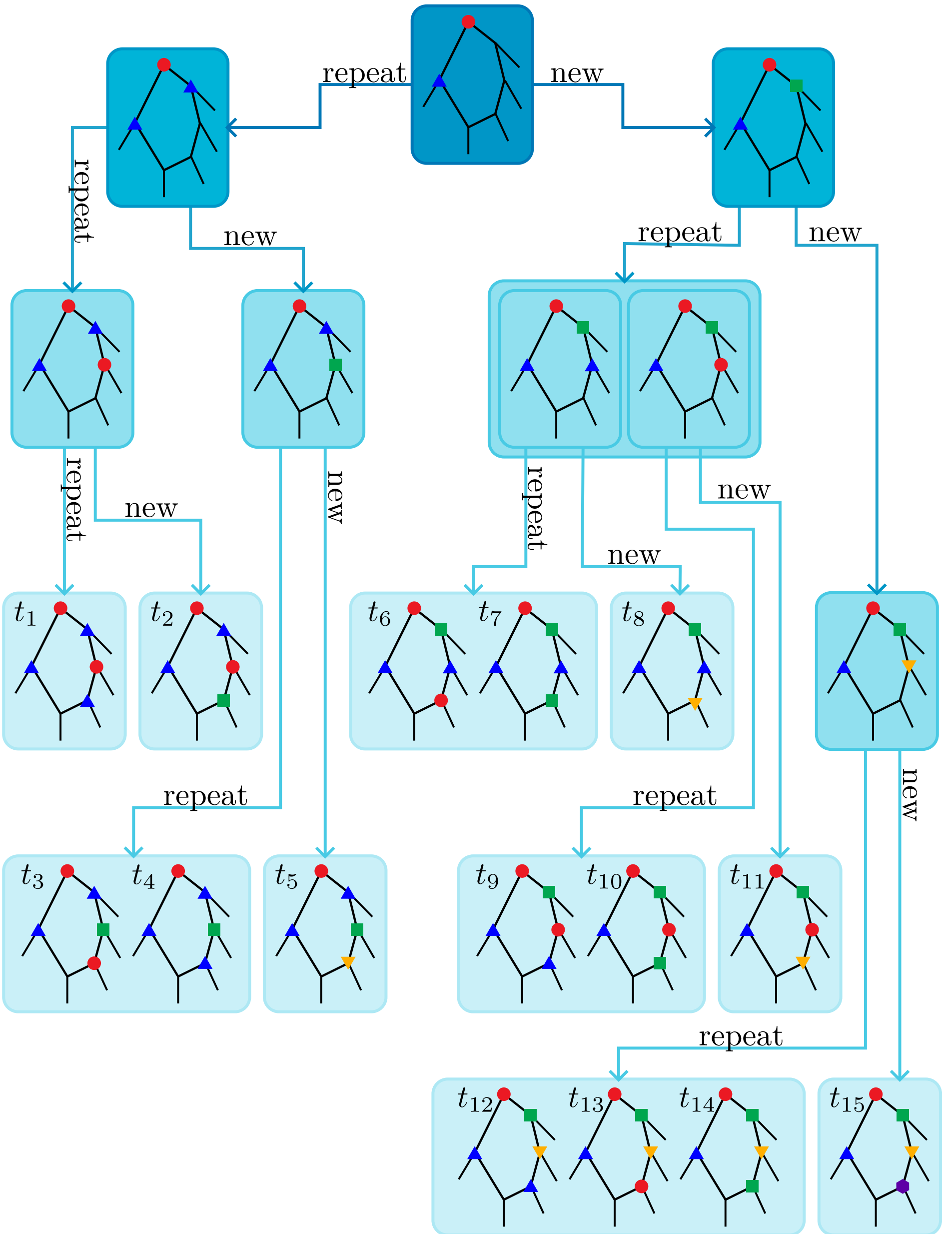


Figure 33: Enumeration of all (▲,●)-fixed quasi-discriminating labelings t_i of $N_{(5)}$. The lighter the bounding box, the fewer vertices are yet to be labeled. In the end, there are 15 distinct t_i, out of which seven have precisely three distinct labels. For readability, the leaves are not labeled but the underlying tree is still $N_{(5)}$ of Figure 31.

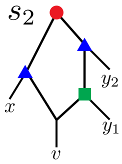
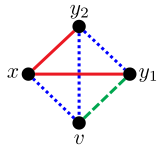
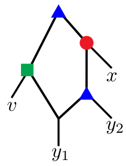
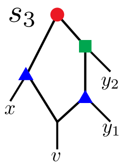
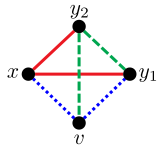
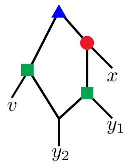
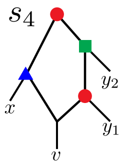
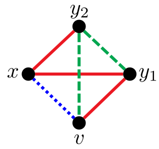
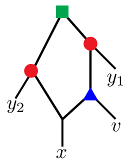
$L_i^{(4)}$	$(N_{(4)}, s_i)$	$\mathcal{G}(N_{(4)}, s_i)$	(N'_i, s'_i)
$L_2^{(4)} = ($▲, ●, ▲, ■$)$			
$L_3^{(4)} = ($▲, ●, ■, ▲$)$			
$L_4^{(4)} = ($▲, ●, ■, ●$)$			

Table 1: Every 4-labeling array $L_i^{(4)}$ with three distinct elements, presented with their corresponding (▲, ●)-fixed labeling s_i, the edge-colored graph $\mathcal{G}(N_{(4)}, s_i)$ and an additional galled tree $(N^{i'}_{(4)}, s'_i)$ such that $\mathcal{G}(N^{i'}_{(4)}, s'_i) = \mathcal{G}(N_{(4)}, s_i)$.

edge-colors, which contradicts condition (C3'). These three and seven edge-colored graphs are presented in Table 1 and Table 2, respectively, alongside their labeling of $N_{(n)}$. For later, we note the following, whose proof can be informally verified by the reader by inspecting Figure 34.

Lemma 4.12. *The edge-colored graphs $\mathcal{G}(N_{(4)}, s_2)$ and $\mathcal{G}(N_{(4)}, s_4)$ are isomorphic, and the edge-colored graphs $\mathcal{G}(N_{(5)}, t_4)$ and $\mathcal{G}(N_{(5)}, t_9)$ are isomorphic.*

Proof. This is a trivial exercise in constructing explicit isomorphisms. Let $V_4 = \{x, v, y_1, y_2\}$ and $V_5 = \{x, v, y_1, y_2, y_3\}$. For $\mathcal{G}(N_{(4)}, s_2) \simeq \mathcal{G}(N_{(4)}, s_4)$ we define $\varphi : \binom{V_4}{2} \to \binom{V_4}{2}$ by the following mappings:

$$\begin{array}{lll} \{x, v\} \mapsto \{y_2, x\} & \{x, y_1\} \mapsto \{y_2, v\} & \{x, y_2\} \mapsto \{y_2, y_1\} \\ \{v, y_1\} \mapsto \{x, v\} & \{v, y_2\} \mapsto \{x, y_1\} & \{y_1, y_2\} \mapsto \{v, y_1\} \end{array}$$

For $\mathcal{G}(N_{(5)}, t_4) \simeq \mathcal{G}(N_{(5)}, t_9)$ we define $\varphi' : \binom{V_5}{2} \to \binom{V_5}{2}$ by:

$$\begin{array}{lll} \{x, v\} \mapsto \{y_3, x\} & \{x, y_1\} \mapsto \{y_3, y_1\} & \{x, y_2\} \mapsto \{y_3, v\} \\ \{x, y_3\} \mapsto \{y_3, y_2\} & \{v, y_1\} \mapsto \{x, y_1\} & \{v, y_2\} \mapsto \{x, v\} \\ \{v, y_3\} \mapsto \{x, y_2\} & \{y_1, y_2\} \mapsto \{y_1, v\} & \{y_1, y_3\} \mapsto \{y_1, y_2\} \\ \{y_2, y_3\} \mapsto \{v, y_2\} & & \end{array}$$

The trivial but tedious verification that φ and φ' are indeed isomorphims of the edge-colored graphs in question is left to the reader. □

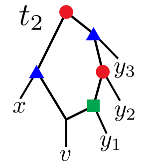
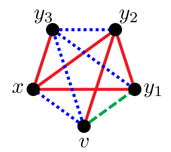
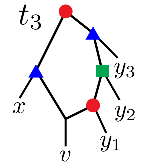
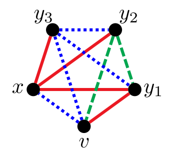
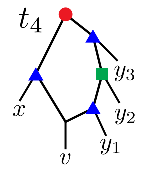
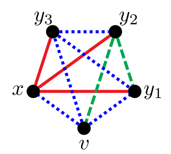
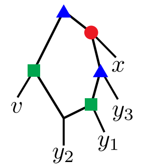
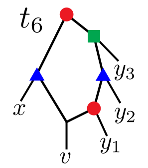
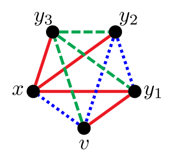
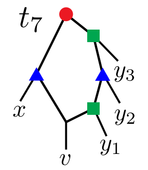
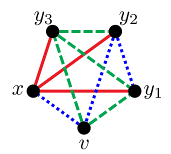
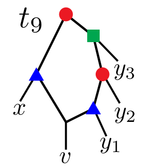
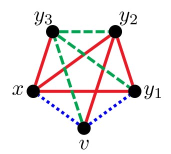
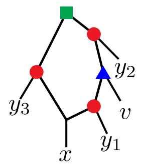
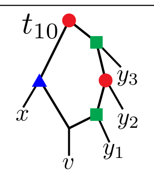
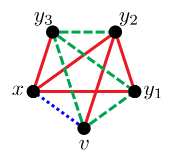
$L_i^{(5)}$	t_i	$\mathcal{G}(N_{(5)}, t_i)$	$(N^{i'}_{(5)}, s'_i)$
$L_2^{(5)} = ({\color{blue}\blacktriangle}, {\color{red}\bullet}, {\color{blue}\blacktriangle}, {\color{red}\bullet}, {\color{green}\blacksquare})$			
$L_3^{(5)} = ({\color{blue}\blacktriangle}, {\color{red}\bullet}, {\color{blue}\blacktriangle}, {\color{green}\blacksquare}, {\color{red}\bullet})$			
$L_4^{(5)} = ({\color{blue}\blacktriangle}, {\color{red}\bullet}, {\color{blue}\blacktriangle}, {\color{green}\blacksquare}, {\color{blue}\blacktriangle})$			
$L_6^{(5)} = ({\color{blue}\blacktriangle}, {\color{red}\bullet}, {\color{green}\blacksquare}, {\color{blue}\blacktriangle}, {\color{red}\bullet})$			
$L_7^{(5)} = ({\color{blue}\blacktriangle}, {\color{red}\bullet}, {\color{green}\blacksquare}, {\color{blue}\blacktriangle}, {\color{green}\blacksquare})$			
$L_9^{(5)} = ({\color{blue}\blacktriangle}, {\color{red}\bullet}, {\color{green}\blacksquare}, {\color{red}\bullet}, {\color{blue}\blacktriangle})$			
$L_{10}^{(5)} = ({\color{blue}\blacktriangle}, {\color{red}\bullet}, {\color{green}\blacksquare}, {\color{red}\bullet}, {\color{green}\blacksquare})$			

Table 2: Every 5-labeling array $L_i^{(5)}$ with three distinct elements only, presented with their respective corresponding $({\color{blue}\blacktriangle}, {\color{red}\bullet})$-fixed labeling t_i, the edge-colored graph $\mathcal{G}(N_{(5)}, t_i)$ and, where possible, an additional galled tree $(N^{i'}_{(4)}, s'_i)$ such that $\mathcal{G}(N^{i'}_{(4)}, s'_i) = \mathcal{G}(N_{(4)}, s_i)$. For proof that none of the gray cells may be filled correctly, see Theorem 4.14

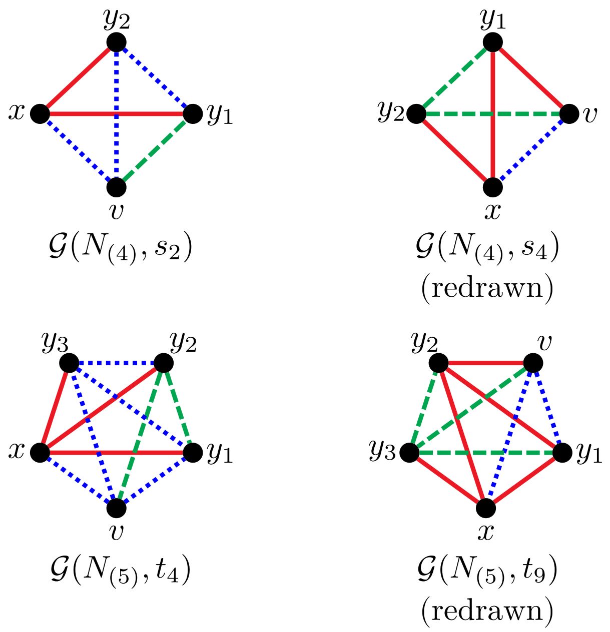


Figure 34: Informal verification that $\mathcal{G}(N_{(4)}, s_2) \simeq \mathcal{G}(N_{(4)}, s_4)$ respectively $\mathcal{G}(N_{(5)}, t_4) \simeq \mathcal{G}(N_{(5)}, t_9)$ indeed holds. Note that the two rightmost graphs are redrawn in comparison to their respective appearance in Table 1 and Table 2.

To characterize polar-cats that are explained by different galled trees we need the following

Definition 4.13. Let $\Sigma_{\triangledown}$ denote a rainbow triangle and put

$$\mathcal{F} := \left\{\Sigma_{\triangledown}, \mathcal{G}(N_{(4)}, s_1), \mathcal{G}(N_{(4)}, s_2), \mathcal{G}(N_{(4)}, s_3), \mathcal{G}(N_{(5)}, t_1), \mathcal{G}(N_{(5)}, t_4)\right\}.$$

Note that $\mathcal{G}(N_{(4)}, s_1) = \Sigma_4$ and $\mathcal{G}(N_{(5)}, t_1) = \Sigma_5$ can be found in Figure 30, $\mathcal{G}(N_{(4)}, s_2)$ and $\mathcal{G}(N_{(4)}, s_3)$ in Table 1 and $\mathcal{G}(N_{(5)}, t_4)$ in Table 2.

Theorem 4.14. *Let Σ be a polar-cat. There are at least two non-isomorphic galled trees that explain Σ if and only if Σ is isomorphic to one of the edge-colored graphs in $\mathcal{F}$.*

Proof. For the "if"-direction it suffices to give two distinct galled trees that explain each of the six polar-cats in $\mathcal{F}$. These are depicted in Figure 27 (rainbow triangle), in Figure 30 ($\mathcal{G}(N_{(4)}, s_1)$ and $\mathcal{G}(N_{(5)}, t_1)$), in Table 1 ($\mathcal{G}(N_{(4)}, s_2)$ and $\mathcal{G}(N_{(4)}, s_3)$) and, lastly, in Table 2 ($\mathcal{G}(N_{(5)}, t_4)$).

For the "only if"-direction, assume that Σ is a polar-cat that is explained by (at least) two distinct galled trees. By Proposition 4.9 this happens only if Σ satisfies condition (C3). If Σ has three vertices only, then it must by Observation 3.1 be (isomorphic to) a rainbow triangle. If it has only two edge-colors then it is isomorphic to $\mathcal{G}(N_{(4)}, s_1)$ or $\mathcal{G}(N_{(5)}, t_1)$. If Σ has at least four vertices and at least three edge-colors in addition to satisfying condition (C3) it, in fact, satisfies condition (C3'). From here on assume so. In particular, Σ contains four or five vertices.

Now, Proposition 4.10, Observation 4.2 and Lemma 4.11 we have $\Sigma \simeq \mathcal{G}(N_{(4)}, s_i)$ for $i \in \{2, 3, 4\}$ if Σ has four vertices, and $\Sigma \simeq \mathcal{G}(N_{(5)}, t_j)$ for $j \in \{2, 3, 4, 6, 7, 9, 10\}$ if Σ has five vertices. Since Lemma 4.12 ensures $\mathcal{G}(N_{(4)}, s_2) \simeq \mathcal{G}(N_{(4)}, s_4)$ we have established the result for polar-cats with four vertices. For five-vertex polar-cats we note that Lemma 4.12 also states that $\mathcal{G}(N_{(5)}, t_4) \simeq \mathcal{G}(N_{(5)}, t_9)$, so it remains only to show that $\mathcal{G}(N_{(5)}, t_i)$ is explained by a unique galled tree for each $i \in \{2, 3, 6, 7, 10\}$.

By Theorem 4.6 it suffices to show that $\mathcal{G}(N_{(5)}, t_i)$ has a fixpoint for each $i \in \{2, 3, 6, 7, 10\}$. To this end we need to establish that, for any fixed $i \in \{2, 3, 6, 7, 10\}$, the graph $\Sigma_i := \mathcal{G}(N_{(5)}, t_i)$ is not a (a, Ω_1, Ω_2)-polar-cat for $a \in \{x, y_1, y_2, y_3\}$. For each $y_j \in \{y_1, y_2, y_3\}$ Proposition 3.8 and Theorem 2.5 implies that it suffices to show that $\Sigma_i - y_j$ is not $\mathfrak{unp}$ by either showing that $\Sigma_i - y_j$ contains a rainbow triangle, or by showing that some monochromatic subgraph of $\Sigma_i - y_j$ is not a cograph i.e. that some monochromatic subgraph of $\Sigma_i - y_j$ contains a P_4. In Table 3, we provide such a rainbow triangle or P_4 in $\Sigma_i - y_j$ for each $i \in \{2, 3, 6, 7, 10\}$ and each $j \in \{1, 2, 3\}$.

As it turns out, $\Sigma_i - x$ is $\mathfrak{unp}$ for each $i \in \{2, 3, 6, 7, 10\}$, so here we instead prove that $\Sigma_i - x$ cannot be a (x, Ω_1, Ω_2)-polar-cat by showing that one of Ω_1 and Ω_2 cannot be $\mathfrak{unp}$ and thus not explained by a caterpillar tree. Note that for each $i \in \{2, 3, 6, 7, 10\}$, we have $\Sigma_i - x = \Sigma_i[\{y_3\}] \otimes_k \Sigma_i[\{v, y_1, y_2\}]$ for some color k. Hence Σ_i may only be a (x, Ω_1, Ω_2)-polar cat if $\{\Omega_1, \Omega_2\} = \{\Sigma_i[\{x, y_3\}], \Sigma_i[\{x, v, y_1, y_2\}]\}$. Table 3 provides a rainbow triangle in $\Sigma_i[\{x, v, y_1, y_2\}]$ or points out that some monochromatic subgraph of $\Sigma_i[\{x, v, y_1, y_2\}]$ contains a P_4, for each $i \in \{2, 3, 6, 7, 10\}$. By Proposition 3.8, this means that either Ω_1 or Ω_2 is not $\mathfrak{unp}$, and Σ_i is not a (x, Ω_1, Ω_2)-polar-cat. We have thus established that Σ_i has a fixpoint for each $i \in \{2, 3, 6, 7, 10\}$, which concludes the proof. □

Theorem 4.6 and Theorem 4.14 can also be combined to directly obtain the following.

Theorem 4.15. *Let Σ be a polar-cat. The following statements are equivalent.*

1. *There is a unique galled tree (N, t) that explains Σ.*

2. *Σ has a fixpoint.*

3. *Σ is not isomorphic to any edge-colored graph in the set $\mathcal{F}$.*

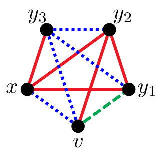
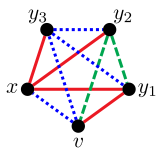
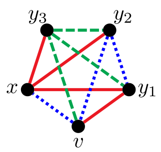
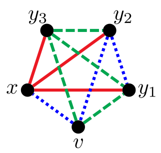
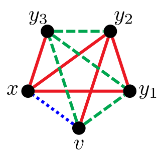
$\Sigma_i^{(5)} = \mathcal{G}(N_{(5)}, t_i)$		Violation
$\Sigma_2^{(5)}$: 	x	$\triangledown(x, v, y_1)$ in $\Omega_i\|_{\bullet}$
	y_1	P_4 in $\left(\Sigma_2^{(5)} - y_1\right)\Big\|_{\bullet}$
	y_2	$\triangledown(x, v, y_1)$ in $\Sigma_2^{(5)} - y_2$
	y_3	$\triangledown(x, v, y_1)$ in $\Sigma_2^{(5)} - y_3$
$\Sigma_3^{(5)}$: 	x	$\triangledown(x, v, y_2)$ in $\Omega_i\|_{\bullet}$
	y_1	$\triangledown(x, v, y_2)$ in $\Sigma_3^{(5)} - y_1$
	y_2	P_4 in $\left(\Sigma_3^{(5)} - y_2\right)\Big\|_{\bullet}$
	y_3	$\triangledown(x, v, y_2)$ in $\Sigma_3^{(5)} - y_3$
$\Sigma_6^{(5)}$: 	x	P_4 in $\Omega_i\|_{\bullet}$
	y_1	$\triangledown(x, v, y_3)$ in $\Sigma_6^{(5)} - y_1$
	y_2	$\triangledown(x, v, y_3)$ in $\Sigma_6^{(5)} - y_2$
	y_3	P_4 in $\left(\Sigma_6^{(5)} - y_3\right)\Big\|_{\bullet}$
$\Sigma_7^{(5)}$: 	x	$\triangledown(x, v, y_1)$ in $\Omega_i\|_{\bullet}$
	y_1	$\triangledown(x, v, y_3)$ in $\Sigma_7^{(5)} - y_1$
	y_2	$\triangledown(x, v, y_1)$ in $\Sigma_7^{(5)} - y_2$
	y_3	$\triangledown(x, v, y_1)$ in $\Sigma_7^{(5)} - y_3$
$\Sigma_{10}^{(5)}$: 	x	$\triangledown(x, v, y_1)$ in Ω_i
	y_1	$\triangledown(x, v, y_3)$ in $\Sigma_{10}^{(5)} - y_1$
	y_2	$\triangledown(x, v, y_3)$ in $\Sigma_{10}^{(5)} - y_2$
	y_3	$\triangledown(x, v, y_1)$ in $\Sigma_{10}^{(5)} - y_3$

Table 3: Collection of, so to speak, 'certificates' on why $\mathcal{G}(N_{(5)}, t_i)$ is not a (a, Ω_1, Ω_2)-polar-cat for $a \in \{x, y_1, y_2, y_3\}$. The notation $\triangledown(a, b, c)$ indicates that the vertices a, b and c induce a rainbow triangle. See the proof of Theorem 4.14 for details.

Recall, by definition, $(N,t) \simeq (N',t')$ implies $\mathcal{G}(N,t) \simeq \mathcal{G}(N',t')$. As seen in Example 2.17, the converse is, in general, not satisfied. Nevertheless, Theorem 4.15 provides conditions under which $\mathcal{G}(N,t) \simeq \mathcal{G}(N',t')$ implies $(N,t) \simeq (N',t')$ for polar-cats $\mathcal{G}(N,t)$ and $\mathcal{G}(N',t')$.

Corollary 4.15.1. *Suppose Σ and Σ' are polar-cats such that there is no $\Sigma'' \in \mathcal{F}$ which satisfies $\Sigma'' \simeq \Sigma$ or $\Sigma'' \simeq \Sigma'$. Let (N,t) and (N',t') be galled trees that explains Σ and Σ', respectively. Then $\Sigma \simeq \Sigma'$ if and only if $(N,t) \simeq (N',t')$.*

With this, we consider the question of when a polar-cat is explained by a unique galled tree to be resolved.

5 Future research

In this thesis we have provided two main characterizations: which primitive edge-colored graphs are explained by at least one galled tree (Theorem 3.2) and which primitive edge-colored graphs are explained by, up to isomorphism, precisely one galled tree (Theorem 4.15). These characterizations are direct generalizations of Theorem 6.9 (see also [21, Thm. 3.7]) respectively Proposition 6.13 of [19], although the wording of the statements and the proof strategies have differed at times. It is natural to believe that further results of [19] are possible to generalize. In particular, it is reasonable that the following conjecture would hold true.

Conjecture 5.1. *An edge-colored graph Σ is explained by a galled tree if and only if $\Sigma[M]/\mathbb{M}_{\max}(\Sigma[M])$ is a polar-cat for every prime module M of Σ. Moreover, it can be verified in polynomial time whether a given edge-colored graph is explained by some galled tree and if so, then a such a galled tree can be constructed in polynomial time.*

To understand the statement above, some terminology needs to be introduced; let $\Sigma = (V, \sigma)$ denote some edge-colored graph. The set $\mathbb{M}_{\max}(\Sigma)$ contains the proper, strong and inclusion-maximal modules of Σ. By definition $\mathbb{M}_{\max}(\Sigma)$ is a partition of V. Note also that if M and M' are disjoint modules Σ, then $\sigma(xy) = \sigma(x'y')$ for all edges $x, x' \in M$ and all $y, y' \in M'$ [8, Lem. 4.11]. We may thus define the *quotient graph* $\Sigma/\mathbb{M}_{\max}(\Sigma)$ as the edge-colored graph with vertex set $V' := \mathbb{M}_{\max}(\Sigma)$ and edge-coloring σ' defined by putting $\sigma'(M, M') := \sigma(x, y)$ for all $M, M' \in V'$, where $x \in M$ and $y \in M'$ are chosen arbitrarily. If (the first statement of) Conjecture 5.1 where to be proved with the same proof-strategy as Theorem 7.5 of [19] the concept of *prime vertex replacement* from [3] would need a counterpart for edge-colored graphs. In other words, even though it is not far-fetched to believe Conjecture 5.1 holds, there will for certain be an extensive workload to prove so.

In extension to (and partly relying on) Conjecture 5.1 it would be of interest to further study the algorithmic aspects of galled trees that explain edge-colored graphs, as has been done for graphs without edge-colors in [20]. There exist a multitude of NP-complete and NP-hard problems for edge-colored graphs that could, potentially, be polynomial time-solvable on graphs $\mathcal{G}(N, t)$ by making use of the labeled galled tree (N, t). One example of this type of difficult problems include finding so-called *colored cuts* i.e. a partition of the vertex set of an edge-colored graph into two sets A and B such that the number of edge-colors of edges between A and B are either minimized or maximized. In practice, colored cuts are useful in, for example, image segmentation: see e.g. [12]. Other interesting NP-complete (respectively NP-hard) problems with varying level of applicability can be found in e.g. [1, 17, 23].

On the more theoretical side of things, there is a previously unmentioned characterization of $\mathfrak{unp}$ edge-colored graphs in [22], which might be possible to generalize to the case of galled trees. The characterization in question states that an edge-colored graph Σ is $\mathfrak{unp}$ if and only if the monochromatic subgraph $\Sigma|_k$ is a cograph for each edge-color k and the so-called *set of 1-clusters* $\mathcal{C}^1(\Sigma)$ of Σ forms a *hierarchy* (i.e. the $\subseteq$-relation on elements of $\mathcal{C}^1(\Sigma)$ is a reflexive and transitive). See also Theorem 2.5. We do not wish to formally define the set $\mathcal{C}^1(\Sigma)$, as it would be a rather technical detour, but note that the condition in essence states that the respective modular decomposition of the monochromatic subgraphs is, so to speak, compatible with the modular decomposition of an edge-colored graph that is explained by a tree. Since the set of 1-clusters of a galled tree was recently characterized rather neatly [18, Thm. 8.9] we see a potential opening in this line of study.

Lastly, we point out that there are two possibilities for even further generalizations than we have managed here: one could work with more general objects than edge-colored graphs or with more general objects than galled trees. Exchanging edge-colored graphs for 2-structures would, for example, be a natural progression of the topic of this thesis. If one would explore how other networks than galled trees can explain edge-colored graphs (or graphs without edge-colors, or 2-structures) there are two important aspects to keep in mind; first and foremost, the network in question must have a unique least common ancestor defined for each pair of leaves (known as lca-networks in [18]). Secondly, if the end-goal of algorithmic applicability is to be retained, then we would expect that the network may not be 'too' complicated. At the very least, one should consider the possibility of utilizing the networks in algorithms in tandem with the work towards characterizations of graphs explained by such networks.

References

[1] A. Abouelaoualim et al. “Paths and trails in edge-colored graphs”. In: *Theoretical Computer Science* 409.3 (2008), pp. 497–510. ISSN: 0304-3975. DOI: 10.1016/j.tcs.2008.09.021.

[2] A. Blass. “Graphs with unique maximal clumpings”. In: *Journal of Graph Theory* 2.1 (1978), pp. 19–24. DOI: 10.1002/jgt.3190020104.

[3] C. Bruckmann, P. Stadler, and M. Hellmuth. “From modular decomposition trees to rooted median graphs”. In: *Discrete Applied Mathematics* 310 (2022), pp. 1–9. ISSN: 0166-218X. DOI: 10.1016/j.dam.2021.12.017.

[4] D. Corneil, H. Lerchs, and L. Stewart Burlingham. “Complement reducible graphs”. In: *Discrete Applied Mathematics* 3.3 (1981), pp. 163–174.

[5] A. Cournier and M. Habib. “A new linear algorithm for modular decomposition”. In: *Trees in Algebra and Programming—CAAP’94: 19th International Colloquium Edinburgh, UK, April 11–13, 1994 Proceedings 19*. Springer. 1994, pp. 68–84.

[6] R. Diestel. *Graph Theory*. 5th Ed. Graduate Texts in Mathematics. Springer-Verlag Berlin Heidelberg, 2017.

[7] A. Ehrenfeucht and G. Rozenberg. “Primitivity is hereditary for 2-structures”. In: *Theoretical Computer Science* 70.3 (1990), pp. 343–358. ISSN: 0304-3975. DOI: 10.1016/0304-3975(90)90131-Z.

[8] A. Ehrenfeucht and G. Rozenberg. “Theory of 2-structures, part I: Clans, basic subclasses, and morphisms”. In: *Theoretical Computer Science* 70.3 (1990), pp. 277–303. ISSN: 0304-3975. DOI: 10.1016/0304-3975(90)90129-6.

[9] A. Ehrenfeucht and G. Rozenberg. “Theory of 2-structures, part II: Representation through labeled tree families”. In: *Theoretical Computer Science* 70.3 (1990), pp. 305–342. ISSN: 0304-3975. DOI: 10.1016/0304-3975(90)90130-A.

[10] A. Ehrenfeucht et al. “An O(n2) Divide-and-Conquer Algorithm for the Prime Tree Decomposition of Two-Structures and Modular Decomposition of Graphs”. In: *Journal of Algorithms* 16.2 (1994), pp. 283–294. ISSN: 0196-6774. DOI: 10.1006/jagm.1994.1013.

[11] J. Engelfriet et al. “Characterization and complexity of uniformly nonprimitive labeled 2-structures”. In: *Theoretical Computer Science* 154.2 (1996), pp. 247–282. ISSN: 0304-3975. DOI: 10.1016/0304-3975(94)00272-X.

[12] L. Faria et al. "On Colored Edge Cuts in Graphs". In: *Anais do I Encontro de Teoria da Computação*. Porto Alegre: SBC, 2016, pp. 780–783. DOI: 10.5753/etc.2016.9764.

[13] P. Gambette, V. Berry, and C. Paul. "Quartets and unrooted phylogenetic networks". In: *Journal of Bioinformatics and Computational Biology* 10.04 (2012). DOI: 10.1142/S0219720012500047.

[14] P. Gambette, K.T. Huber, and S. Kelk. "On the challenge of reconstructing level-1 phylogenetic networks from triplets and clusters". In: *J. Math. Biol.* 74 (2017), pp. 1729–1751. DOI: 10.1007/s00285-016-1068-3.

[15] M. Habib and C. Paul. "A survey of the algorithmic aspects of modular decomposition". In: *Computer Science Review* 4.1 (2010), pp. 41–59. ISSN: 1574-0137. DOI: 10.1016/j.cosrev.2010.01.001.

[16] P. Halmos. *Naive Set Theory*. 1st Ed. Undergraduate Texts in Mathematics. Springer New York, 1998. DOI: 10.1007/978-1-4757-1645-0.

[17] P. Hell, Y. Manoussakis, and Z. Tuza. "Packing problems in edge-colored graphs". In: *Discrete Applied Mathematics* 52.3 (1994), pp. 295–306. ISSN: 0166-218X. DOI: 10.1016/0166-218X(94)90148-1.

[18] M. Hellmuth, D. Schaller, and P. Stadler. *Clustering Systems of Phylogenetic Networks*. 2022. arXiv: 2204.13466 [q-bio.PE].

[19] M. Hellmuth and GE. Scholz. "From modular decomposition trees to level-1 networks: Pseudo-cographs, polar-cats and prime polar-cats". In: *Discrete Applied Mathematics* 321 (2022), pp. 179–219. ISSN: 0166-218X. DOI: 10.1016/j.dam.2022.06.042.

[20] M. Hellmuth and GE. Scholz. *Linear Time Algorithms for NP-hard Problems restricted to GaTEx Graphs*. 2023. arXiv: 2306.04367 [cs.DM].

[21] M. Hellmuth and GE. Scholz. *Resolving Prime Modules: The Structure of Pseudo-cographs and Galled-Tree Explainable Graphs*. 2022. arXiv: 2211.16854 [math.CO].

[22] M. Hellmuth, P. Stadler, and N. Wieseke. "The mathematics of xenology: di-cographs, symbolic ultrametrics, 2-structures and tree-representable systems of binary relations." In: *Journal of mathematical biology* 1.75 (2017), pp. 199–237. DOI: 10.1007/s00285-016-1084-3.

[23] Z. Huang and X. Li. "Hardness results for three kinds of colored connections of graphs". In: *Theoretical Computer Science* 841 (2020), pp. 27–38. ISSN: 0304-3975. DOI: 10.1016/j.tcs.2020.06.030.

[24] D. Huson, R. Rupp, and C. Scornavacca. *Phylogenetic Networks: Concepts, Algorithms and Applications*. Cambridge University Press, 2010. DOI: 10.1017/CBO9780511974076.

[25] A. Lindeberg. *Generate k-labeling arrays.* https://github.com/AnnaLindeberg/kLabelingArrays. July 19, 2023.

[26] R. McConnell and J. Spinrad. "Modular decomposition and transitive orientation". In: *Discrete Mathematics* 201.1 (1999), pp. 189–241. ISSN: 0012-365X. DOI: 10.1016/S0012-365X(98)00319-7.

[27] R. Möhring. "Algorithmic Aspects of Comparability Graphs and Interval Graphs". In: *Graphs and Order: The Role of Graphs in the Theory of Ordered Sets and Its Applications.* Ed. by I. Rival. Springer Netherlands, 1985, pp. 41–101. ISBN: 978-94-009-5315-4. DOI: 10.1007/978-94-009-5315-4_2.

[28] L. N. Shevrin and N. D. Filippov. "Partially ordered sets and their comparability graphs". In: *Siberian Mathematical Journal* 11 (1970), pp. 497–509. DOI: 10.1007/BF00967091.

[29] M. Tedder et al. "Simpler Linear-Time Modular Decomposition Via Recursive Factorizing Permutations". In: *Automata, Languages and Programming.* Ed. by L. Aceto et al. Springer Berlin Heidelberg, 2008, pp. 634–645. ISBN: 978-3-540-70575-8.

List of corrections

In Section 2.3

On page 14, the sentence "As noted in [19] condition (N4) implies that ..." should instead say "As noted in [19] condition (N3) and (N4) implies that ...".

In proof of Lemma 3.7

The last sentence ("Since C was arbitrarily chosen, we thus conclude that N has a single cycle C such that $\rho_C = \rho_N$.") in the first paragraph of the proof on page 42 is a faulty conclusion. It may, correctly, be replaced with the following two sentences:

"Since C was arbitrarily chosen, we thus conclude that any cycle C of N satisfies $\rho_C = \rho_N$. Moreover, by condition (N4) in the definition of a galled tree, any two cycles of N are edge-disjoint."

In the paragraph that follows, C is defined to be "the cycle" (of N), which then must be replaced with "a cycle" (of N). The remainder of the proof and in particular the argument that $L^C(\rho_N) = \emptyset$ is nevertheless still valid. In particular, $L^C(\rho_N) = \emptyset$ and $|L^C(v)| = 1$ for each $v \in V^0(C) \setminus \{\rho_N\}$ implies that there is no cycle in N other than C.

In proof of Proposition 4.8

On page 55, the following sentence appears after equation (18):

"For this to be true at the same time as (18), we must have that $v, z \in V(\Pi_i)$ for either $i = 1$ or $i = 2$, while $y \in V(\Pi_j)$ for $j \in \{1, 2\} \setminus \{i\}$."

However, the second part of the sentence is not necessarily true; possibly $y \in V(\Pi_i)$. Nevertheless, whether $y \in V(\Pi_1)$ or $y \in V(\Pi_2)$, we still have that $v, z, x \in V(\Pi_i)$ for $i = 1$ or $i = 2$, hence Π_i contains the rainbow triangle $\Sigma[\{x, v, z\}]$, and the rest of the proof is correct as-is.