Gaps in the Categories of Finite Directed and Finite Transitive Graphs

Robert Baldinger Jonathan McCammond

AMS Mathematics Subject Classification (1990 revision):
05C20, 18A10, 68R10

0 Introduction

The problem of classifying the gaps in the Tilson ordering for the categories of finite directed and finite transitive graphs was first proposed by John Rhodes in [5]. In the present article we give a complete classification of the gaps in these categories. More specifically, given an arbitrary finite directed or finite transitive graph G, we give here a necessary and sufficient condition on the structure of G for there to exist a gap $H \prec G$. In addition, a procedure is given for constructing the graph H whenever the graph G satisfies the necesssary conditions. Many of the methods used here are similar to those used in [2]. Many of these results have also been shown in [1].

The structure of the article is as follows. The first section contains the basic definitions and key examples used in the rest of the article. In the second section, the gaps in the category of finite directed graphs are classified, and in the third section the classification is extended to the category of finite transitive graphs. The fourth and final section relates these results to the previous work in the field, and it concludes with the statement of an open problem.

1 Basic Definitions and Examples

In this first section we collect the basic definitions and examples used in the remainder of the article. Notice that we have followed the convention of writing all function and relation symbols to the right of their arguments.

1.1 Directed and Transitive Graphs

A directed graph G can be compactly described as a pair of disjoint sets O and A together with a pair of functions D and C from A to O. Directed graphs are sometimes referred to as digraphs and will usually be designated by upper-case Roman letters such as G. The elements of O are called objects, the elements of A are called arrows, and the functions D and C assign each arrow to its domain and codomain respectively. Objects will be denoted by lower-case Roman letters, and arrows by lower-case Greek letters. Thus if $G = (O, A, D, C)$ is a directed graph, then $a \in O$ is a typical object of G and $\alpha \in A$ is a typical arrow. Graphically, an arrow α with domain a and codomain b will be written $a \xrightarrow{\alpha} b$, or even $a \to b$ if α is a unique arrow from a to b. The set of all objects and arrows of G will also be referred to as $\mathrm{Obj}(G)$ and $\mathrm{Arr}(G)$, respectively. If both $\mathrm{Obj}(G)$ and $\mathrm{Arr}(G)$ are finite sets, then G is called a finite directed graph. Since all of the graphs in this article will be finite and directed, the words 'graph', 'digraph' and 'finite directed graph' will henceforth be used synonymously. Also note that in practice, graphs are only distinguished up to isomorphism (defined below). Thus Saunders MacLane[3, p. 7] gives $\bullet \to \bullet \to \bullet$ and $\bullet \rightrightarrows \bullet$ as examples of finite digraphs. Statements of the form 'G is the unique digraph such that' some property is true are generally interpreted to mean that G is the unique digraph, up to isomorphism, for which that property holds.

The set of all arrows in G from a to b is called a hom-set of G and it is denoted $G(a, b)$. Notice that $\mathrm{Arr}(G)$ is a disjoint union of the hom-sets of G. A graph G will be called a transitive graph iff $G(a, b) \neq \emptyset$ and $G(b, c) \neq \emptyset$ always implies $G(a, c) \neq \emptyset$. An arrow which starts and ends at the same object is called a local arrow, and a graph with no local arrows is called locally trivial. If for all distinct objects a and b either $G(a, b)$ or $G(b, a)$ is empty, then G is known as a 1-way graph. Let a and b be objects in a graph G. If $G(a, b)$ is not empty, then a is called a predecessor of b, and b is a successor of a. If a and b are also distinct, then a is a proper predecessor of b, and b is a proper successor of a. The objects a in G which have a proper predecessor and a proper successor will be called interior objects. An object with a proper successor but no proper predecessor is called a source object. An object with a proper predecessor but no proper successor is called a sink object. And an object with neither a proper predecessor nor a proper successor is called an isolated object.

A subgraph H of G is a graph $H = (O', A', D', C')$ such that $O' \subseteq O$, $A' \subseteq A$, with $D' = D \mid A'$ and $C' = C \mid A'$. A full subgraph of G is a subgraph H such that $H(a, b) = G(a, b)$ for all objects $a, b \in \mathrm{Obj}(H)$. Notice that full subgraphs are completely determined by the objects they contain, so that $G(O')$, where O' is a subset of $\mathrm{Obj}(G)$, is an unambiguous notation for the full subgraph of G determined by the objects in O'. For distinct objects a and b in a graph G, let $G\{a, b\} = G(a, b) \cup G(b, a)$. If $G\{a, b\}$ is nonempty, then a and b are said to be linked. The objects a and b are called connected in G if there is a sequence of vertices $c_0, c_1, \ldots, c_n$ with $c_0 = a$, $c_n = b$ and c_i linked with c_{i+1} for all $i = 0, 1, \ldots, n - 1$. The integer n is called the length of the connection. Connections of length zero are allowed, in which case $a = b = c_0$. As usual, the graph G is called connected if and only if every pair of objects in G is connected, and every graph G can be uniquely decomposed into a disjoint union of connected components which are full subgraphs of G.

The following common examples of graphs will be used in the course of the proof. The unique graph with no objects and no arrows is called the empty graph and is denoted by $\boldsymbol{\emptyset}$. For each integer $n > 0$, $\boldsymbol{n}$ represents the digraph with n objects labeled $0, 1, \ldots, n-1$ and no arrows. The digraph $\mathbf{1}$ is also called $\bullet$ following MacLane's method of specifying digraphs, alluded to above. Next, for any integer $n > 0$, $\boldsymbol{n}^\bullet$ represents the graph obtained from $\boldsymbol{n}$ by adding exactly one arrow for each object of $\boldsymbol{n}$ and by requiring that each added arrow start and end at its associated object. The graph $\vec{\boldsymbol{n}}$, $n \geq 0$, is the graph (O, A, D, C) with objects O labeled by the set $\{0, 1, \ldots, n\}$, arrows A labeled by the set $\{(i, j) : 0 \leq i < j \leq n\}$, and with D and C defined as the projections of A onto the first and second coordinate of the labels, respectively. At the other extreme there are directed graphs with only a single object. For each integer $n \geq 0$, let L_n be the digraph with one object denoted $*$, and n arrows which must, of course, start and end at $*$.

As a final example, let $s_1, s_2, \ldots, s_l$ be a finite sequence of integers with $l > 0$. The graph $E_{s_1, s_2, \ldots, s_l}$ is the digraph with $l+1$ objects labeled $0, 1, \ldots, l$, such that for each i from 0 to l there are s_i arrows from i to $i+1$ when s_i is positive, and there are $-s_i$ arrows from $i+1$ to i when s_i is negative. The examples from MacLane given above are $E_{1,1}$ and E_2, respectively, in this notation.

1.2 Morphisms and Divisions between Graphs

Let G and H be directed graphs. A morphism Φ from H to G, denoted $\Phi : H \to G$, can be thought of as a family of functions with various domains. Since each of these functions can be distinguished by their domains, they will all be denoted Φ. The specific functions constituting Φ include an object function Φ from $\mathrm{Obj}(G)$ to $\mathrm{Obj}(H)$ and, for each pair of objects a and b in $\mathrm{Obj}(G)$, a hom-set function Φ from $G(a, b)$ to $H(a\Phi, b\Phi)$. A morphism is called faithful if each of its hom-set functions is injective. An embedding is a faithful morphism with an injective object function. An isomorphism is a morphism whose object function, and all hom-set functions, are bijections. Clearly, for every graph G there is an identity morphism from G to itself which is an isomorphism, and for every subgraph H in G there is an embedding of H into G. The category of all finite directed graphs and their morphisms will be denoted **FDG**, and the category of all finite transitive graphs and morphisms will be denoted **FTG**.

The following examples are illustrative. Given $n > m > 0$ there is an embedding of E_m into E_n and a non-faithful morphism from E_n to E_m. There are exactly three morphisms from E_1 to $E_{1,-1,1}$, each one of which is an embedding, and there is only one morphism from $E_{1,-1,1}$ to E_1, and it is faithful. There are exactly two morphisms from E_2 to $E_{1,1}$ but no morphisms from $E_{1,1}$ to E_2. Finally, notice that there is a faithful morphism from G to L_n, where n is the size of the largest hom-set of G.

A division Φ from H to G, denoted $\Phi : H \precsim G$, can be thought of as an object function and an arrow relation. As above, the letter Φ can be used for all of these relations without risk of confusion. Specifically the function and the relation are an object function Φ from the objects of H to the objects of G, and a relation Φ between the arrows of H and the arrows of G such that for all arrows α from a to b in H, $\alpha\Phi$ is a nonempty subset of $G(a\Phi, b\Phi)$, and for distinct arrows α and β in $H(a,b)$, $\alpha\Phi \cap \beta\Phi = \emptyset$. Notice that a morphism is a division if and only if it is faithful, and that, in fact, the following three statements are equivalent: (1) $H \precsim G$, (2) there exists a faithful morphism $\Phi : H \to G$, and (3) there is a function $f : \mathrm{Obj}(H) \to \mathrm{Obj}(G)$ such that for all objects a, b of H, $|H(a,b)| \leq |G(af, bf)|$. Notice that $H \precsim G$ does not imply $|\mathrm{Obj}(H)| \leq |\mathrm{Obj}(G)|$, as is illustrated by the faithful morphism from $E_{1,-1,1}$ to E_1.

A graph H is called equivalent to a graph G if and only if H divides G and G divides H, and H strictly divides G if and only if H divides G but G does not divide H. These situations will be denoted $H \sim G$ and $H \prec G$, respectively. A gap from H to G occurs when $H \prec G$ and every K with $H \precsim K \precsim G$ has either $K \sim H$ or $K \sim G$. It is easy to see that the relation $\precsim$ is reflexive and transitive on **FDG**, and that the quotient category of **FDG** by the relation $\sim$ under $\precsim$ is a countably infinite upper semilattice whose least upper bound is the join of two graphs. Moreover, the join $\bigvee_{i \in I} G_i \precsim G$ if and only if each $G_i \precsim G$. Proofs of these facts can be found in [6].

1.3 Constructions on Graphs

Let $(G_i)_{i \in I}$ be an indexed family of graphs, where the index set I is finite. The join $\bigvee_{i \in I} G_i$ is the disjoint union of the indexed graphs. That is, if each $G_i = (O_i, A_i, D_i, C_i)$, then the join is the graph (O, A, D, C), where O is the disjoint union of the sets O_i, A is the disjoint union of the sets A_i, and D and C are the unique functions from A to O such that $D \,|\, A_i = D_i$ and $C \,|\, A_i = C_i$. We write $G_1 \vee \cdots \vee G_n$ for $\bigvee_{i \in \{1,\ldots,n\}} G_i$. Similarly, the product $\prod_{i \in I} G_i$ is the direct product of the indexed graphs. That is, if each $G_i = (O_i, A_i, D_i, C_i)$, the product is the graph (O, A, D, C), where O is the direct product of the sets O_i, A is the direct product of the sets A_i, and D and C are the unique functions from A to O such that the projection onto the i-th coordinate A_i followed by D_i (or C_i) equals the function D (or C) followed by the projection onto O_i. We write $G_1 \times \cdots \times G_n$ for $\prod_{i \in \{1,\ldots,n\}} G_i$.

One way of constructing a new graph from a given graph G is to start with the direct product $K = G \times \boldsymbol{n}^\bullet$. The objects and arrows of K are ordered pairs whose first coordinate represents an object or an arrow of G and whose second coordinate represents its level. Projection onto the first coordinate is a faithful morphism from K to G, and for all i the full subgraph of objects with second coordinate i forms an isomorphic copy of G called G_i. See Figure 1. When new arrows are added to the graph K, the added arrows will be called global arrows to distinguish them from the 'local' arrows already in K. Notice that every local arrow may be labeled by an ordered triple of the form $(\alpha, (a,i), (b,i))$ where α

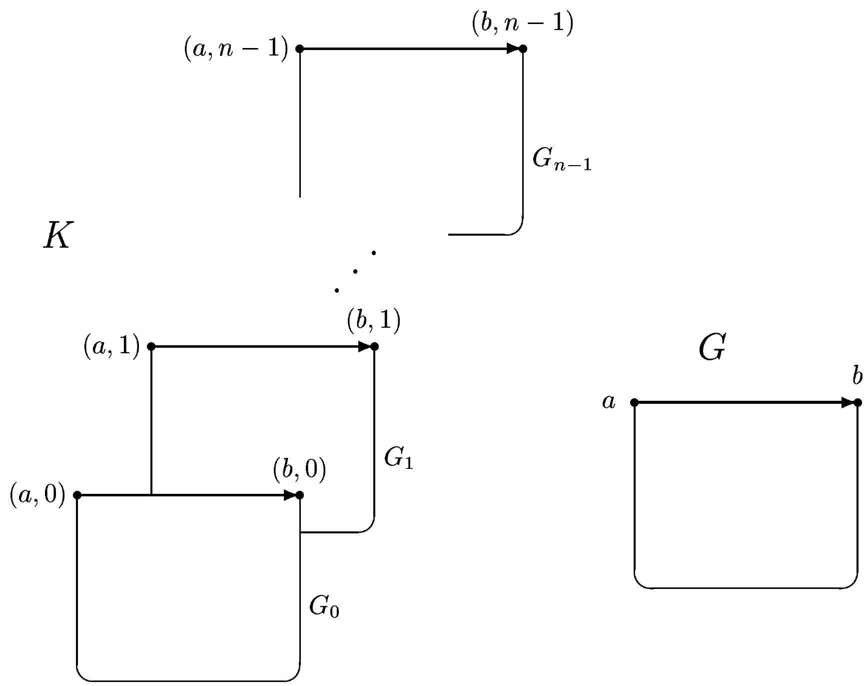

Figure 1: Graph K constructed from n copies of G. A nonempty hom-set $G(a, b)$ is portrayed here by a single bold arrow.

is an arrow in G from a to b, and that every global arrow has a label of the form $(\alpha, (a, i), (b, j))$, where i and j may or may not be identical and α may or may not be an arrow in G.

1.4 Reductions of Graphs

Let H be a full subgraph of G. If there is a division $\Phi : G \precsim H$ such that the restriction of Φ to H is the identity morphism on H, then H is called a retract of G and Φ is called a retraction from G to H. If H is a retract of G and no proper full subgraph of H is a retract of G, then H is called the reduction of G. A graph whose only retract is itself is known as a reduced graph. Almost by definition, a reduction of a graph G is reduced and equivalent to G, a reduction of a reduction of G is a reduction of G, and G is reduced if and only if G is isomorphic to one of its reductions. The notion of a 'reduction' of a graph, category, or other directed structure was introduced by Rhodes in [6]. By the following lemma, we may speak of the unique reduction of G, up to isomorphism.

Lemma 1.1 *If G and H are reduced graphs, then G and H are equivalent if and only if they are isomorphic. In particular, any two reductions of a graph G are isomorphic.*

Notice that if G is reduced then every component of G must be reduced, and the distinct components must be pairwise incomparable in the $\precsim$ ordering.

Lemma 1.2 *If G and H are graphs which are reduced and equivalent, then every division $\Phi : H \precsim G$ is an isomorphism. In particular, if G is reduced then every division $\Phi : G \precsim G$ is an automorphism of G.*

Proof: Since G and H are equivalent there exists a division $\Psi : G \precsim H$. Notice first that the objects in $H\Phi$ must be all of $\mathrm{Obj}(G)$ or else the division $\Psi\Phi : G \precsim G$ shows that G is equivalent to one of its proper full subgraphs, contradicting the assumption that G is reduced. Thus, since all of the sets involved are finite, and since by Lemma 1.1, $\mathrm{Obj}(H)$ and $\mathrm{Obj}(G)$ have the same size, $\Psi\Phi$ merely permutes the objects of G. Set $k = |\mathrm{Obj}(G)|!$, so that the division $(\Psi\Phi)^k$ is the identity on the objects of G. Next, since each hom-set is sent to itself, the relation on arrows must actually be a function, so that $(\Psi\Phi)^k$ is a faithful morphism from G to G, and this morphism at worst permutes the arrows. If n is the size of the largest hom-set of G, let $l = k(n!)$, so that $(\Psi\Phi)^l$ is now the identity morphism on all of G. A similar argument shows that $(\Phi\Psi)^l$ is the identity morphism on H, so that $(\Psi\Phi)^{l-1}\Psi$ is a two-sided inverse of Φ. It follows quickly that Φ is bijective on the arrows in each hom-set, completing the proof of the first statement. The second is an immediate consequence of the first. □

Let $\Phi : H \precsim G$ be a division and let K be a full subgraph of H. If $\Phi\,|\,K : K \precsim G$ is an isomorphism, then K is called a Φ-readable copy of G in H.

Lemma 1.3 (Labeling Lemma) *If $H \sim G$ with G reduced, then for every division $\Phi : H \precsim G$, there exists a Φ-readable copy of G in H. In particular, if G is reduced and $\Phi : H \precsim G$ is a division whose object function is not onto, then H strictly divides G.*

Proof: Let R be the reduction of H, so that R is reduced and equivalent to H and G. By Lemma 1.1 R is isomorphic to G, and by Lemma 1.2 the division $\Phi\,|\,R : R \precsim G$ is an isomorphism producing a Φ-readable copy of G in H. □

1.5 Gaps between Graphs

In an earlier article [2] we showed that, up to equivalence, the only gaps between finite, connected, undirected graphs are

$$\emptyset \prec \bullet \prec E_1 \prec E_2 \prec E_3 \prec \ldots$$

where the E_n denotes the undirected graph containing exactly 2 vertices and connected to each other by n undirected edges.[pp We will conclude this introductory section by showing that there are connected gaps between finite directed graphs corresponding to each of the connected gaps between finite undirected graphs, and by showing that all gaps between directed graphs must be locally trivial. The results below follow the approach used by Rhodes in [6].

Lemma 1.4 *If K strictly divides E_n with $n > 0$, then K divides E_{n-1}. This shows that the divisions $\bullet \prec E_1 \prec E_2 \prec \ldots$ are gaps in* **FDG**. *Trivially, $\emptyset \prec \bullet$ is a gap in* **FDG**.

Proof: If K strictly divides E_n, then there exists a faithful morphism $\Psi : K \to E_n$ and no faithful morphism from E_n to K. The latter fact implies that any hom-set of K sent to $E(0,1)$ is not surjective, and this in turn shows that there is a faithful morphism from K to E_{n-1} with the same object function. Thus $E_{n-1} \prec E_n$ is a gap. The final statement of the lemma is clear. □

Lemma 1.5 *A graph G divides a graph H if and only if $G \times \vec{\boldsymbol{n}}$ divides H for all n.*

Proof: If G divides H, then it is clear by projection onto the first coordinate that $G \times \vec{\boldsymbol{n}} \precsim G \precsim H$ for all n, so suppose that $G \times \vec{\boldsymbol{n}}$ divides H for all n, and fix n so that $n > |\mathrm{Obj}(H)|^{|\mathrm{Obj}(G)|}$. For this n let $K = G \times \vec{\boldsymbol{n}}$, let $\Phi : K \to H$ be a faithful morphism arising from the assumed division, and let Φ_i be the function from $\mathrm{Obj}(G)$ to $\mathrm{Obj}(H)$ defined by $g\Phi_i = (g,i)\Phi$. Since there are only $|\mathrm{Obj}(H)|^{|\mathrm{Obj}(G)|}$ such functions, there must exist distinct integers i and j such that $\Phi_i = \Phi_j$. Next, consider arbitrary objects a and b in G. Since

$$|G(a,b)| = |K((a,i),(b,j))| \leq |H(a\Phi_i, b\Phi_j)| = |H(a\Phi_i, b\Phi_i)|$$

it follows that Φ_i is the object function of a faithful morphism from G to H. Thus G divides H and the proof is complete. □

Lemma 1.6 *If G is a connected graph which is not locally trivial, then $H \prec G$ is not a gap. In particular, there is a finite directed graph K with $H \prec K \prec G$.*

Proof: It is immediate that $H \precsim H \vee (G \times \vec{\boldsymbol{n}}) \precsim G$ for each n, and since G is connected, any faithful morphism from G to $H \vee (G \times \vec{\boldsymbol{n}})$ must send G into either a connected component of H or into $G \times \vec{\boldsymbol{n}}$. By assumption and by the fact that G is not locally trivial, neither of these options are viable possibilities, so that $H \precsim H \vee (G \times \vec{\boldsymbol{n}}) \prec G$. Moreover, by Lemma 1.5 there is an integer n such that $G \times \vec{\boldsymbol{n}}$ does not divide H. Setting $K = H \vee (G \times \vec{\boldsymbol{n}})$ for this n completes the proof. □

2 Gaps in FDG

In this section we investigate the gaps in the category of finite directed graphs. We will begin by focusing specifically on connected graphs. For this special case, we narrow down the types of divisions which can give rise to gaps, and then we show by construction that all of the remaining possibilities do indeed produce gaps. In the final portion of the section we extend these results to all of the graphs in **FDG**. As a final note, we should mention that the wordings used in the proofs in this section have often been chosen so that they will apply without alteration to the circumstances considered in Section 3.

2.1 Connected Graphs

The following lemmas will show, among other results, that if G is a connected graph and $H \prec G$ is a gap in **FDG**, then it can be assumed without loss of generality that H is also connected. First we will show a special case of this result.

Lemma 2.1 *Let G be a reduced, connected, and locally trivial graph, let H_1 and H_2 be connected graphs, and let $\Phi_1 : H_1 \prec G$, and $\Phi_2 : H_2 \prec G$ be strict divisions. If the subgraphs $H_1\Phi_1$ and $H_2\Phi_2$ in G are not disjoint, then there exists a connected graph K such that H_1 and H_2 divide K and K strictly divides G.*

Proof: We may assume without loss of generality that Φ_1 and Φ_2 are faithful morphisms and that $H_1\Phi_1$ and $H_2\Phi_2$ have an object in common, since when two subgraphs have an arrow in common they must necessarily have the endpoints of the arrow in common as well. Fix $a_1 \in \mathrm{Obj}(H_1)$ and $a_2 \in \mathrm{Obj}(H_2)$ with $a_1\Phi_1 = a_2\Phi_2$. Call this common image a. If one of the objects a_i, say a_1, has neither a proper predecessor nor a proper successor, then since G is locally trivial it follows that H_1 is the graph $\bullet$ and that the lemma is true with K equal to H_2. Thus we may assume that both a_1 and a_2 have either a proper predecessor or a proper successor. For concreteness, suppose that a_1 has a proper predecessor b_1 in H_1 and that a_2 has a proper successor b_2 in H_2. The other possibilities can be treated analogously.

A new graph K can be constructed from $H_1 \vee H_2$ by adding three new objects c_1, d, and c_2, and four new arrows. In the particular case assumed here there are arrows from c_1 to a_1, from c_1 to d, from d to c_2, and from a_2 to c_2. See Figure 2. Clearly, K is connected and each H_i divides K. Moreover, K divides G since there is a faithful morphism Φ which agrees with Φ_1 on the objects of H_1 and with Φ_2 on the objects of H_2 and which sends c_1 to $b_1\Phi_1$, d to a, and c_2 to $b_2\Phi_2$. If K were equivalent to G, then by the Labeling Lemma (Lemma 1.3) there would exist a Φ-readable copy G' of G in K. Since G is connected, and since a_1, a_2, and d are all sent to the object a in G, it follows that G' is contained in the full subgraph on $\mathrm{Obj}(H_1) \cup \{c_1\}$, the full subgraph

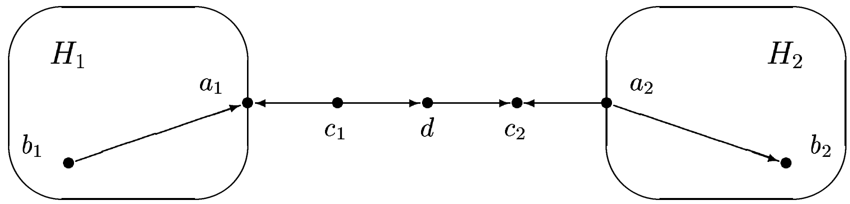

Figure 2: Graph K constructed in the proof of Lemma 2.1

on $\mathrm{Obj}(H_2) \cup \{c_2\}$, or the full subgraph on $\{c_1, d, c_2\}$. Since the first two cases are easily seen to be equivalent to H_1 and H_2, respectively, contradicting the assumption that Φ_1 and Φ_2 are strict divisions, the only possibility is that G' and thus G contains at most three objects and two edges. The lemma is easily seen to be true for the severely restricted possibilities which remain. $\Box$

Lemma 2.2 *If G is connected and locally trivial, then the family of connected graphs strictly below G is upward-directed in the sense that if H_1 and H_2 are connected and $\Phi_1 : H_1 \prec G$ and $\Phi_2 : H_2 \prec G$ are strict divisions, then there is a connected K such that H_1 and H_2 divides K and K strictly divides G.*

Proof: As in the previous lemma we may assume without loss of generality that G is reduced, that each Φ_i is a faithful morphism, and that neither H_1 nor H_2 is the empty graph or the graph $\bullet$. In particular, since G is locally trivial we may assume that H_1 and H_2 and their images in G contain at least two distinct objects. By Lemma 2.1 we may further assume that the subgraphs $H_1\Phi_1$ and $H_2\Phi_2$ in G are disjoint. Since G is connected, there is a connection from $H_1\Phi_1$ to $H_2\Phi_2$, say objects $c_0, c_1, \ldots, c_n$ in G, where $c_0 = a_1\Phi_1$, $c_n = a_2\Phi_2$, and $G\{c_i, c_{i+1}\}$ is nonempty for all $i = 0, \ldots, n-1$. If the connection chosen is the shortest possible between $H_1\Phi_1$ and $H_2\Phi_2$, then the objects $c_1, c_2, \ldots, c_{n-1}$ will be disjoint from $H_1\Phi_1 \vee H_2\Phi_2$ in G.

Let H_1' be a graph constructed from H_1 by adding exactly n new objects, $b_1, b_2, \ldots, b_n$, and exactly n new arrows, one each between b_i and b_{i+1} for $i = 0, 1, \ldots, n-1$ whose orientation depends on the orientation of the arrows in $G\{c_i, c_{i+1}\}$. It is clear that H_1 divides H_1' and that H_1' divides G by a faithful morphism whose object function agrees with Φ_1 on the objects of H_1 and sends b_i to c_i. Since $H_2\Phi$ has more than one object, and since the objects $c_1, c_2, \ldots c_{n-1}$, $H_1\Phi_1$, and $H_2\Phi_2$ are pairwise disjoint, it follows that the image of H_1' does not include all of the objects of G and thus by the Labeling Lemma, H_1' strictly divides G. Lemma 2.1 applied to H_1' and H_2 completes the proof. $\Box$

Lemma 2.3 *If G is connected and $H_1, H_2 \prec G$ are gaps in* **FDG***, then $H_1 \sim H_2$, and there is a graph $K \sim H_1 \sim H_2$ with K connected.*

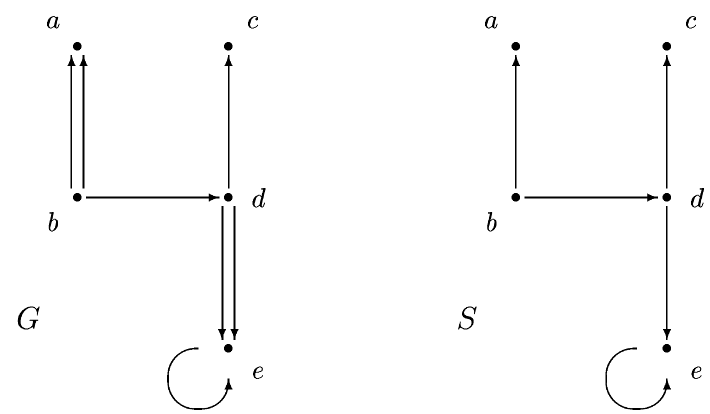

Figure 3: The skeleton S of a graph G

Proof: First, G is locally trivial, by Lemma 1.6. Therefore Lemma 2.2 can be applied repeatedly to the finite number of components of $H_1 \vee H_2$ to yield a connected K divided by each component of $H_1 \vee H_2$ and yet still strictly below G. Since $H_1, H_2 \prec G$ are gaps, we must have $H_1 \sim K \sim H_2$. $\square$

2.2 Skeletons of Graphs

We will now show that the structure of an arbitrary reduced, connected graph G which is part of a gap $H \prec G$ in **FDG** must have a 'skeleton' which is a 'tree'. We will begin with some definitions relating to the shape of a graph. The skeleton of a graph G is a subgraph S with the same objects as G, but S contains exactly one arrow from each nonempty hom-set of G. Such a subgraph is clearly unique up to isomorphism, and notice also that G is connected if and only if its skeleton is connected. The skeleton of a graph is analogous to the graph $G_D(C)$ associated to a category C as defined by Rhodes in [6]. Finally, notice that if the skeleton of G is reduced, then this implies that G itself is reduced.

A hom-set of a graph G is said to split G, if G is connected but the subgraph obtained by removing the arrows in this hom-set is disconnected. Notice that a graph G is a tree in the traditional sense if and only if G is isomorphic to its skeleton and every nonempty hom-set splits G. More generally, if the skeleton of G is a tree, then G is connected and locally trivial. Figure 3 illustrates these definitions. In the figure, S is not a tree because the hom-set $S(e, e)$ does not split S. If G' is the graph obtained from G by deleting the local arrow at e, then the skeleton of G' will be a tree.

Let a and b be objects of some graph G, and let α be a new arrow from a to b disjoint from the existing arrows in $G(a, b)$. The graph G with the arrow α added will be variously denoted as $G + (a \xrightarrow{\alpha} b)$, $G + (a \to b)$, or simply G^α. More generally, we call any graph with the same objects but a finite number of additional arrows an extension of the original graph. Next, let $\Phi : H \to G$ be a faithful morphism. A hom-set $H(a,b)$ of H is called Φ-full if $|H(a,b)| = |G(a\Phi, b\Phi)|$. A graph H is called Φ-saturated if for every hom-set $H(a,b)$ which is not Φ-full, adding a new arrow to $H(a,b)$ produces a strictly larger graph in the sense that $H \prec H + (a \to b)$. Finally, let $H \prec G$ be a gap where H is Φ-saturated. A hom-set $G(c,d)$ of G is called Φ-critical if $G(c,d)$ splits G and $G(c,d)$ has a non-full pre-image in H, that is, if there exist $a \in c\Phi^{-1}$ and $b \in d\Phi^{-1}$ such that $H(a,b)$ is not Φ-full. The next three results are technical lemmas which will be used in the proof of Lemma 2.7.

Lemma 2.4 *If $\Phi : H \to G$ is a faithful morphism, then there is a graph H' which contains H as a subgraph and a faithful morphism $\Phi' : H' \to G$ which agrees with Φ on H, such that H' has the same objects as H, H' is equivalent to H, and H' is Φ'-saturated.*

Proof: If H is not Φ-saturated, then by definition there is a hom-set $H(a,b)$ which is not Φ-full and H is equivalent to $H + (a \to b)$. In this case extend the faithful morphism Φ to the new arrow in the obvious way and repeat this procedure as many times as necessary. At some point the process must stop since the total number of arrows in any extension H' of H which divides G is bounded above by the number of hom-sets in H times the size of the largest hom-set in G. When it does stop, the graph and the faithful morphism under consideration satisfy the conclusion of the lemma. $\square$

Lemma 2.5 (Splitting Lemma) *Let G be a connected and reduced graph, let $H \prec G$ be a gap in* **FDG**, *and let $\Phi : H \to G$ be a faithful morphism with a surjective object function. If H is Φ-saturated and $H(a,b)$ is a hom-set which is not Φ-full, then its image under Φ, $G(a\Phi, b\Phi)$, is Φ-critical.*

Proof: Let H' be the graph $H + (a \to b)$. Since H is Φ-saturated and $H(a,b)$ is not Φ-full, $H \prec H' \precsim G$. The fact that $H \prec G$ is a gap now implies that H' and G are equivalent. Let $N = |H'(a,b)|$ so that $|H(a,b)| = N - 1$. We now fix some $n > |\mathrm{Obj}(H)|$, and construct a graph K from $H \times \boldsymbol{n}^{\bullet}$ by adding N new arrows connecting (a,i) to (b,j) for all integers with $0 < i < j < n$. See Figure 4. The levels of K, called H_i, are isomorphic copies of H so that H clearly embeds in K. On the other hand, the object function which sends (c,i) in K to $c\Phi$ is the object function of a faithful morphism so that $H \precsim K \precsim G$. Since $H \prec G$ is a gap, K must be equivalent to either H or G.

If K is equivalent to H, then there exists a faithful morphism $\Psi : K \to H$, and since $n > |\mathrm{Obj}(H)|$, there exist distinct integers i and j with $0 \le i < j < n$ and $(b,i)\Psi = (b,j)\Psi$. This implies that the size of the hom-set $H((a,i)\Psi, (b,i)\Psi) = H((a,i)\Psi, (b,j)\Psi)$ is at least N, and that there is a faithful morphism from H' to H obtained by sending the objects c in H' to $(c,i)\Psi$. This contradicts the fact that H strictly divides H' and shows that K is not equivalent to H and that K must be equivalent to G.

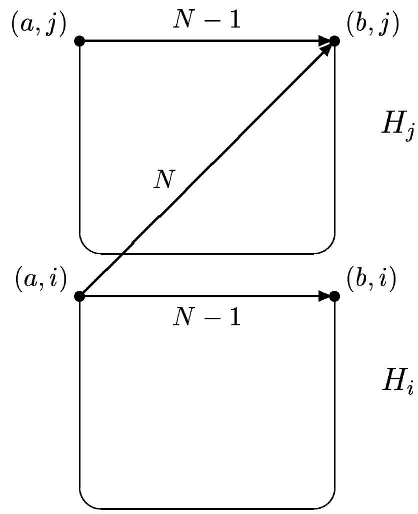

Figure 4: A detail in the construction of K from n copies of H

Now that we know that K is equivalent to G it follows by Lemma 1.3 that there is a Φ'-readable copy G' of G in K. In this context this implies that G' contains exactly one object with first coordinate a and one object with first coordinate b. If G' contains objects (a,i) and (b,i) for some i, then it cannot contain any of the objects (a,j) or (b,j) with $j \neq i$, and in particular, it cannot contain any of the global hom-sets of K. Since G is connected, it follows that G divides the i-th level of K. But since this level is isomorphic to H, we have reached a contradiction. Thus G' must contain (a,i) and (b,j) with $i \neq j$. Since G' cannot contain any of the other objects with first coordinate a or b, G' contains at most one global hom-set, namely, the hom-set $K((a,i),(b,j))$. If i is greater than j, then this hom-set is empty, contradicting the connected nature of G and G'. On the other hand, if i is less than j, then since G' is isomorphic to G, it is clear that $G'((a,i),(b,j))$ splits G' and thus $G(a,b)$ splits G. $\square$

Lemma 2.6 (Continued-splitting) *Let G be a connected and reduced graph, let $H \prec G$ be a gap in* **FDG**, *let $\Phi : H \to G$ be a faithful morphism with a surjective object function, and let H be Φ-saturated. Suppose further that $G(a,b)$ is Φ-critical, and that there is an object c in G such that one of the hom-sets $G(a,c)$, $G(c,a)$, $G(b,c)$, or $G(c,b)$ is non-empty. If Σ represents this non-empty hom-set, then there is a graph K which is equivalent to H and contains H as a full subgraph, and there is a faithful morphism $\Psi : K \to G$, such that K is Ψ-saturated and the hom-set Σ in K is Ψ-critical.*

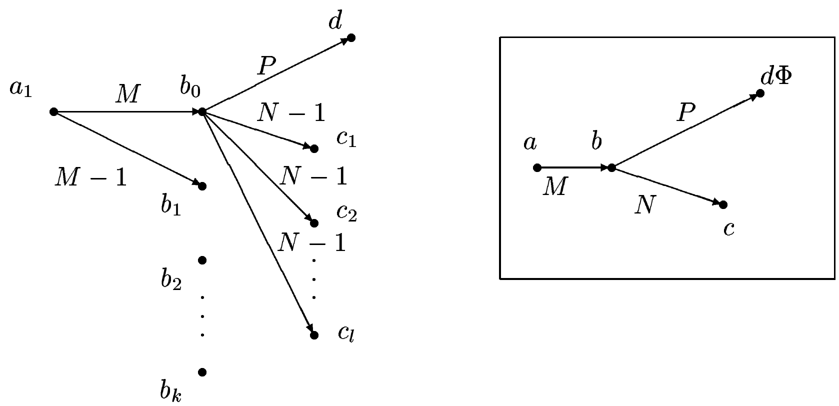

Figure 5: A detail of the construction of H' from H in the proof of Lemma 2.6. The boxed inset shows the corresponding hom-sets in G.

Proof: For the sake of concreteness, let $\Sigma = G(b,c)$. The other cases are treated similarly. Let $M = |G(a,b)|$ and $N = |G(b,c)|$. Next, choose $a_1 \in a\Phi^{-1}$ and $b_1 \in b\Phi^{-1}$ so that $|H(a_1,b_1)| = M-1$. Also let $b\Phi^{-1} = \{b_1,\ldots,b_k\}$ and $c\Phi^{-1} = \{c_1,\ldots,c_l\}$ with $k,l > 0$ be a listing of all of the inverse images of b and c. We will construct a new graph H' from H as follows. First, add a single new object called b_0 and add $N-1$ arrows from b_0 to each of the objects $c_1,\ldots,c_l$. Then add $|G(b,d\Phi)|$ new arrows from b_0 to d and $|G(d\Phi,b)|$ new arrows from d to b_0 for each object $d \in \mathrm{Obj}(H) \setminus \{c_1,\ldots,c_l\}$. In particular, this involves adding M new arrows from a_1 to b_0. See Figure 5. Clearly H is a full subgraph of H' and also there is a faithful morphism $\Phi' : H' \to G$ which agrees with Φ on H and sends b_0 to b, so that $H \precsim H' \precsim G$. Since $H \prec G$ is a gap, H' is equivalent either to G or to H.

If H' were equivalent to G, then by Lemma 1.3 there would be a Φ'-readable copy G' of G in H'. Since $G \not\precsim H$, G' would have to contain b_0 as well as c_i for some i, but since $H'(b_0,c_i)$ is not full, this is impossible. Thus H' cannot be equivalent to G and so it must be equivalent to H instead.

Next, by Lemma 2.4, we can replace the graph H' with an extension H'', and the morphism Φ' with a faithful morphism $\Phi'' : H'' \to G$, such that ϕ'' agrees with Φ' on H', H'' has the same objects as H', $H'' \sim H' \sim H$, and H'' is Φ''-saturated. We now wish to show that not all of the hom-sets $H''(b_0,c_i)$ are Φ''-full. First notice that since H is already Φ-saturated, adding a single new arrow from a_1 to b_1, by Lemma 1.3, produces a Φ-readable copy G' of G in $H + (a_1 \to b_1)$. The full subgraph G' must, of course, include a_1, b_1, and c_i for some i, and since it contains b_1 it cannot contain b_0. Next, suppose that

$H''(b_0, c_i)$ is Φ''-full. Then there is an embedding of G' in the full subgraph of H'' on the objects $(\mathrm{Obj}(G') \setminus \{b_1\}) \cup \{b_0\}$, which is given by fixing all of the objects in $\mathrm{Obj}(G') \setminus \{b_1\}$ and sending b_1 to b_0. Since this would result in an embedding of G in $H'' \sim H$, the hom-set $H''(b_0, c_i)$ must not be Φ''-full and by Lemma 2.5, $G(b, c)$ is Φ''-critical. Setting $K = H''$ and $\Psi = \Phi''$ completes the proof. $\square$

Lemma 2.7 *If G is a reduced and connected graph and $H \prec G$ is a gap in* **FDG**, *then G is locally trivial and the skeleton of G is a tree.*

Proof: We will first show that the lemma is true if we assume that there exists a faithful morphism $\Phi : H \to G$ with a surjective object function. By Lemma 2.4, we may assume without loss of generality that H is Φ-saturated. Next, there must exist a hom-set $H(a, b)$ which is not Φ-full, since otherwise we can select arbitrary pre-images of each object in G (which is possible since Φ is surjective on objects), and thus we can find a full subgraph of H isomorphic to G, contradicting the assumed strictness of the division $H \prec G$. By Lemma 2.5, the image of this non-full hom-set, $G(a\Phi, b\Phi)$, is Φ-critical.

Call an object c in G a type-a object if there is a connection between c and $a\Phi$ in the skeleton of G which does not pass through $b\Phi$, and a type-b object if there is a connection between c and $b\Phi$ in the skeleton of G which does not pass through $a\Phi$. Since the skeleton of G is connected, every object of G is either type-a or type-b but not both, since the hom-set $G(a\Phi, b\Phi)$ splits G. Suppose that $G(c, d)$ is a hom-set corresponding to a hom-set in the skeleton of G, and that c and d are both type-a objects. We must show that $G(c, d)$ splits G. There is a connection $e_0, e_1, \ldots, e_n$ where $e_0 = a\Phi$, either c or d is e_{n-1} and the other is e_n. Moreover, if this is the smallest such connection, it follows that all of the objects e_i are also of type-a. By inductively applying Lemma 2.6, each hom-set $G(e_{i-1}, e_i)$, or $G(e_i, e_{i-1})$ as appropriate, is Φ_i-critical for some graph H_i containing H and some faithful morphism $\Phi_i : H_i \to G$, and therefore this hom-set splits G. This shows that the skeleton of G must be a tree, and by Lemma 1.6 it is already known that G must be locally trivial.

It only remains to show that our initial assumption of a morphism with surjective object function was justified. First, we may assume that H is non-empty, for if $H = \emptyset$, then G must be the graph $\bullet$, and the lemma is clearly true in this case. Next, let $\Phi : H \to G$ be a faithful morphism, and let $a_1, \ldots, a_n$ be the objects of G which are not in the image of H under Φ. We then form a new graph H' from H by adding new objects $b_1, \ldots, b_n$ and extending Φ to $\Phi' : H' \to G$ by sending b_i to a_i for all $i = 1, \ldots, n$. Since H is not empty, H' is equivalent to H, and $H' \sim H \prec G$ is still a gap. This completes the proof. $\square$

2.3 Gaps between Connected Graphs

At this point we have shown that whenever we are given a connected graph G and a gap $H \prec G$, we can assume without loss of generality that H is connected, that G is reduced and locally trivial, and that the skeleton of G is a tree. We will now show by construction that given any G which is a reduced, connected, locally trivial graph whose skeleton is a tree, there exists another graph H, unique up to equivalence, such that $H \prec G$ is a gap. This converse of Lemma 2.7 will complete the classification of connected gaps in the category of finite directed graphs.

Let G be a reduced, connected, locally trivial, and nonempty graph whose skeleton S is a tree. First we need to establish some notation. Fix some object $r \in \mathrm{Obj}(S) = \mathrm{Obj}(G)$ called the root of S. For distinct $a, b \in \mathrm{Obj}(S)$, a will be called an ancestor of b and b is a descendant of a if every connection in S between r and b passes through a. The object a is called a parent of b and b is a child of a if b is a descendant of a and the link $S\{a, b\}$ is nonempty. We will write $\#a$ for the number of children of a. A leaf is an object with no children. Since $G = \bullet$ is an easy case to analyze, we may assume, without loss of generality, that leaves are distinct from the root. The set of descendants of a together with the object a itself, will be called the legacy of a. Next, for each $a \in \mathrm{Obj}(S)$, we will fix some arbitrary ordering of the children of a, and write $a^1, a^2, \ldots, a^{\#a}$ for the first, second, $\ldots$, last child of a respectively.

The construction involves making finitely many copies of each object of S, and then creating links between the copies of an object and the copies of its children. A step-by-step illustration of the entire construction can be found in Figure 6. The i-th copy of an object a will be denoted a_i, and the k-th copy of the j-th child of a will be denoted a^j_k. In general, whenever a link $H\{a_i, a^j_k\}$ is nonempty, the orientation of the nonempty hom-set will match the orientation of the nonempty hom-set in the link $G\{a, a^j\}$. The link $H\{a_i, a^j_k\}$ is called full if $|H\{a_i, a^j_k\}| = |G\{a, a^j\}|$, and almost-full if $|H\{a_i, a^j_k\}| = |G\{a, a^j\}| - 1$. All of the nonempty links in the construction will be either full or almost-full. Notice that a link $G\{a, a^j\}$ cannot be empty, but that an almost-full link in H may be empty.

For each $a \in \mathrm{Obj}(S)$, a locally trivial graph H_a will now be recursively constructed, starting with the leaves of S. The object set of the graph H_a will consist of $(1 + \#b)$ copies of each b in the legacy of a. When a is a leaf, then H_a is the graph $\bullet$, the graph with one object, labeled a_1, and no arrows. Next, assume that a is not a leaf, and that the construction has already been completed from each of its children. That is, assume that H_{a^i} is already defined for each $i = 1, 2, \ldots, \#a$. The graph H_a begins with the graph $\bigvee_{i=1}^{\#a} H_{a^i}$ and $1 + \#a$ new objects labeled $a_1, a_2, \ldots, a_{1+\#a}$. For each $i = 1, \ldots, \#a$, add a non-full link between the ith copy of a, and the last copy of the ith child of a. In notation, $H_a\{a_i, a^i_{1+\#a^i}\}$ is non-full. A full link is added to all of the other possible links between a copy of a and a copy of any of the vertices in the legacy of a. Notice that since G is locally trivial and its skeleton is a tree, this is equivalent to adding a full link to all of the other possible links between a copy of a and a copy of a child of a. The result is the graph H_a. Once H_r has been constructed, define H to be the full subgraph of H_r on the objects of H_r excluding $r_{1+\#r}$.

Lemma 2.8 *Let G be an arbitrary reduced, connected, locally trivial, and non-empty graph whose skeleton S is a tree. If H_r, and H are the directed graphs constructed above, then H_r and G are equivalent, and H strictly divides G.*

Proof: The object function Φ which sends all of the copies of a in H_r to the a in G used in its label shows that H_r divides G, so it only remains to show that G divides H_r. On the other hand, by construction the full subgraph of H_r on the objects $a_{1+\#a}$ is an isomorphic copy of G. Thus H_r and G are equivalent.

Next, let $\Phi : H_r \to G$ be the faithful morphism described above, let G' be a Φ-readable copy of G in H which must exist by Lemma 1.3, and let $\Psi : G \to G'$ be the unique isomorphism from G to G' for which $\Phi\Psi$ is the identity on G. We will show by induction that each $a\Psi = a_{1+\#a}$ for all a in G. This is certainly true for the leaves of G since there is only one copy of each leaf in H_r. Next, assume that this has been shown for each child b of a and consider $a\Psi$. Since the link from a_i to $a^i_{1+\#a^i}$ is only almost full for each $i = 1, 2, \ldots, \#a$, the only possible image of a under Ψ is $a_{1+\#a}$. This completes the induction, and shows, among other things, that there is only one isomorphic copy of G in H_r. In particular, the full subgraph H, since it does not contain the object $r_{1+\#r}$, does not contain an isomorphic copy of G, and thus it strictly divides G. $\square$

Lemma 2.9 *Let G be an arbitrary reduced, connected, locally trivial, and non-empty graph whose skeleton S is a tree and let H be the graph constructed above. If K is any graph such that K strictly divides G then K divides H. This immediately implies that $H \prec G$ is a gap in* **FDG**.

Proof: Let $\Phi : K \to G$ be a faithful morphism. We inductively define a faithful morphism $\Psi : K \to H_r$ as follows. If a is a leaf of G then send all of the objects of K in $a\Phi^{-1}$ to the unique leaf a_1 in H_r. Next, suppose that the object function of Ψ has been defined on $b\Phi^{-1}$ for all of the descendents b of a, and define Ψ on $a\Phi^{-1}$ as follows. Given an object c in $a\Phi^{-1}$ define $c\Psi$ to be the copy of a in H_r with the smallest possible subscript. In particular $c\Psi$ will be $a_{1+\#a}$ if and only if there are objects c^i in K, $i = 1, \ldots, \#a$, which have been sent to the last copy of the child a^i in H_r, and each of the links between c and c^i are Φ-full. In notation, c^i is sent to $a^i_{1+\#a^i}$ and $|K(c, c^i)| = |G(a, a^i)|$. This completes the inductive definition of Ψ. Notice that if there is an object in K which is sent to $r_{1+\#r}$, then by repeatedly using the statement given above, we can work out from the root to the leaves and find objects in K which form an isomorphic copy of G in K. In particular, this shows that if K strictly divides G, then since it does not contain such a copy of G, the definition of Ψ given above, defines a faithful morphism from K into H, completing the proof of the first statement. The second follows immediately from the first and Lemma 2.8. $\square$

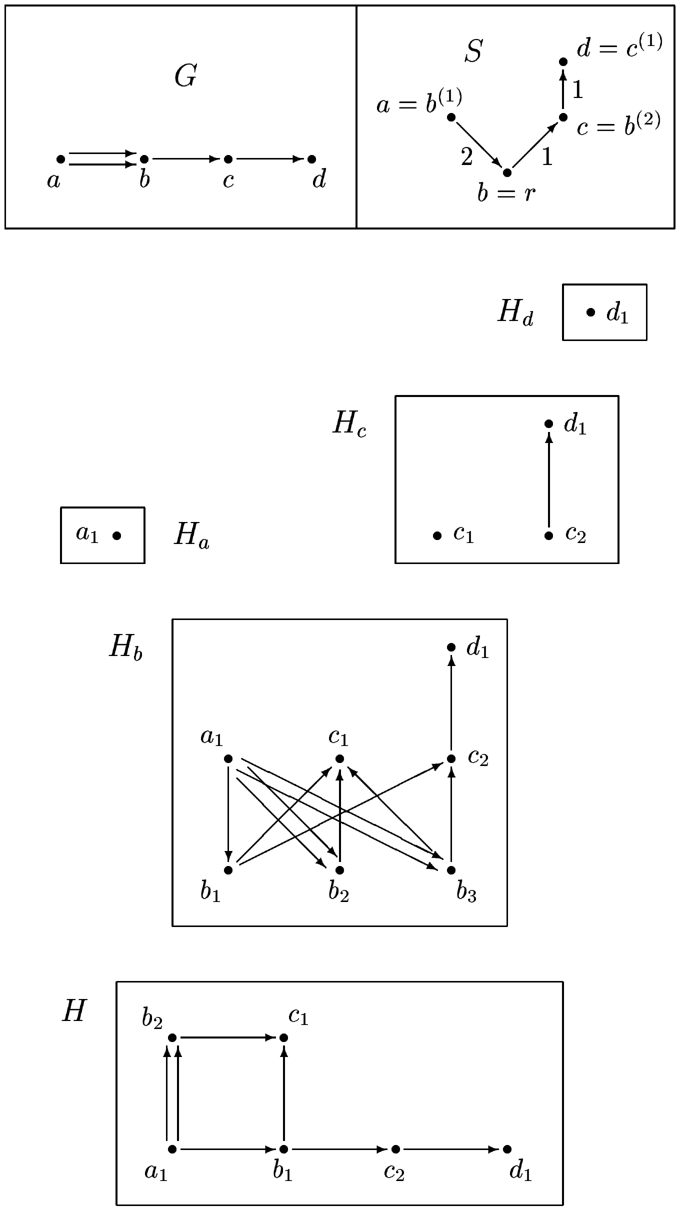

Figure 6: Sample construction of a gap in **FDG**. $G \cong E_{2,1,1}$ is shown in the upper left box. The object b is arbitrarily chosen as the root in S (upper right box). Any other choice of root leads to a divisionally equivalent result for H.

Theorem 2.10 *If G is a nonempty, connected, and reduced directed graph, then there is a gap $H \prec G$ in* **FDG** *if and only if the skeleton of G is a tree. Moreover, when such a gap exists, the directed graph H is unique up to equivalence.*

Proof: The theorem is a combination of Lemmas 2.3, 2.7, and 2.9. □

Recall that the gaps specified in Lemma 1.4 exactly correspond to all the connected gaps in **FG**. But since there are trees in **FDG** which do not reduce to directed graph E_i, that list does not include all gaps between connected finite digraphs. In general gaps in **FDG** have a much richer structure than gaps in **FG**. First, even though given G there is a unique graph H up to equivalence for which $H \prec G$ is a gap, the reverse is false in **FDG**. For example, it is now easy to show that $E_1 \prec E_2$ and $E_1 \prec E_{1,1}$ are both gaps in **FDG** even though E_2 and $E_{1,1}$ are not equivalent. Next, there exist infinite descending sequences of gaps in **FDG**. As an example let $s(n)$ be the finite sequence $(1,1)(-1,1)^{n-1}(1)$, where juxtaposition of finite sequences denotes concatenation and exponentiation denotes multiple concatenations. It is not difficult to verify that each graph $E_{s(n)}$ is reduced, and that the divisions $\cdots \prec E_{s(3)} \prec E_{s(2)} \prec E_{s(1)} = E_{1,1,1}$ are gaps in **FDG**.

2.4 Gaps between Disconnected Graphs

Having classified the gaps in the subcategory of finite connected digraphs, it is now possible to extend this classification to the category of all finite directed graphs. As in the connected case, we may assume without loss of generality that G is reduced. The first two lemmas will show that the disconnected case can be reduced to the connected one. First a definition: a graph H is called close to G if the reduction of G has n connected components and exactly $(n-1)$ of these components divide the graph H.

Lemma 2.11 *If H strictly divides G, but H is not close to G, then $H \prec G$ is not a gap in* **FDG**.

Proof: First, if all n components of G divide H, then G itself divides H and the division is not strict. Next, assume that G_1 and G_2 are two distinct connected components of the reduction of G which do not divide H, and consider the graph $K = H \vee G_1$. Since G_1 and G_2 are distinct connected components of the reduction, it follows that neither graph divides the other. There are obvious divisions $H \precsim K \precsim G$. Clearly, H strictly divides K since G_1 does not divide H, and just as clearly K strictly divides G since G_2 does not divide H or G_1. Thus $H \prec G$ is not a gap. □

Lemma 2.12 *Given graphs $H \precsim K \precsim G$ with G reduced and H close to G, then there exist graphs H', K', G', and L such that G' is connected, $H' \precsim K' \precsim G'$, and $H \sim H' \vee L$, $K \sim K' \vee L$, and $G \sim G' \vee L$.*

Proof: By Lemma 2.11, we may assume that H is close to G. Define L to be the join of the $(n-1)$ connected components of G which divide H and let G' be the unique connected component which does not divide H. Clearly $L \vee G' \sim G$. Next, since L divides H, H is equivalent to $H \vee L$, and, by reduction, H is also equivalent to $H' \vee L$ where H consists only of those connected components of H which do not divide L. Notice that the possibility that H' is the empty graph is allowed. A similar analysis can be applied to K since L also divides K via H. Thus there is a possibly empty graph K' such that $K \sim K' \vee L$ and no connected component of K' divides L. Finally, since $H' \vee L \precsim K' \vee L \precsim G' \vee L$, but no component of H' or K' divides L it must be the case that $H' \precsim K' \precsim G'$. $\square$

Theorem 2.13 *If G is reduced then there exists a graph H for which $H \prec G$ is a gap in* **FDG** *if and only if G contains a connected component G' which is locally trivial and whose skeleton is a tree.*

Proof: By Lemmma 2.12 the sufficiency of the stated conditions is immediate. More specifically, suppose that G' is a component as described and let L be the union of the other components. By the earlier construction there is a graph H' such that $H' \prec G'$ is a gap. Define $H = H' \vee L$. Since G' does not divide L or H' it follows that H strictly divides G and by Lemma 2.12 it is a gap. Thus it only remains to show that these conditions are necessary. Let $H \prec G$ be a gap. Since by Lemma 2.11 we may assume that H is close to G, let G' be the unique connected component of G which does not divide H and let L be the union of the other components. As above let H be written in the form $H' \vee L$ where H' does not divide L. All of the earlier arguments can now be repeated with the graph L attached and with a few minor modifications. Since the details are very nearly identical, we will give only a single illustrative example. Recall that by Lemma 1.5 there is an n sufficiently large such that $G' \times \vec{\boldsymbol{n}}$ does not divide H. If G' were not locally trivial, then, as in the proof of Lemma 1.6, we could create a graph $H' \vee (G' \times \vec{\boldsymbol{n}}) \vee L$ which is strictly between $H = H' \vee L$ and $G = G' \vee L$, thus showing that G' must be locally trivial. A similar alteration of the proofs of the other lemmas succeeds in showing that the skeleton of G' must be a tree. $\square$

Notice that unlike the connected case, when G is disconnected, there may be more than one H for which $H \prec G$ is a gap. For example, if $G = E_2 \vee E_{1,1}$ then H equal to either E_2 or $E_{1,1}$ forms a gap $H \prec G$. In fact the connected components of G which are locally trivial and whose skeletons are trees are in one-to-one correspondence with the gaps H for which $H \prec G$ is a gap. This completes our analysis of gaps in **FDG**.

3 Gaps in FTG

In this section we shift our focus from the category of all finite directed graphs (**FDG**) to the subcategory of all finite transitive graphs (**FTG**). Since the transitivity of the graphs involved is irrelevant to the existence of a division, the $\precsim$-ordering on **FTG** is a restriction of the $\precsim$-ordering on **FDG**. In particular, given transitive graphs G and H, H divides G, H strictly divides G, or H is equivalent to G in **FTG** if and only if the respective statements are true when G and H are viewed as graphs in **FDG**. The existence of a gap on the other hand, which is dependent on the existence or non-existence of a third graph can and does change under this restriction. Certainly gaps $H \prec G$ in **FDG** where both H and G are transitive will also be gaps in **FTG**, so that, for example, the divisions $\emptyset \prec \bullet \prec E_1 \prec E_2 \ldots$ are still gaps in **FTG**, but new gaps are also created. As an example of the latter, consider $\vec{\mathbf{1}} \prec \vec{\mathbf{2}}$. By Theorem 2.10 this division is not a gap in **FDG**, but it will be shown below to be a gap in **FTG**.

More generally, all of the arguments given in Section 1 are unaffected by the restriction to transitive graphs, and thus they remain valid in **FTG**. In particular, since full subgraphs of transitive graphs are automatically transitive, the reduction of a transitive graph is still well defined and transitive (Lemma 1.1), since the Labeling Lemma never uses transitivity, whenever G is reduced, $H \sim G$, and $\Phi : H \to G$ is a faithful morphism, there is a Φ-readable copy of G in H (Lemma 1.3), and since $G \times \vec{\boldsymbol{n}}$ is transitive whenever G is transitive, the transitive graphs G for which $H \prec G$ is a gap must be locally trivial (Lemma 1.6).

3.1 Connected Transitive Graphs

As stated earlier, a transitive graph is simply a directed graph G which is transitive, so that $a \to b \to c$ in G implies that the homset $G(a, c)$ is nonempty. Notice that although a full subgraph of a transitive graph must be transitive, an arbitrary subgraph need not be. For any directed graph, however, there does exist a 'smallest' transitive graph which contains it as a subgraph. If G is a directed graph, the transitive closure of G is the graph $\overline{G}$, obtained from G by adding a single new arrow $a \to b$ to every empty hom-set $G(a, b)$ for which there exist arrows

$$a \to c_1 \to c_2 \to \cdots \to c_n \to b$$

in G. The transitive graph $\overline{G}$ is the smallest transitive graph containing G in the sense that if H is a transitive graph which contains G as a subgraph, it must also contain $\overline{G}$ as a subgraph.

Lemma 3.1 *Let G be a reduced, connected, and locally trivial transitive graph, let H_1 and H_2 be connected transitive graphs, and let $\Phi_1 : H_1 \prec G$, and $\Phi_2 : H_2 \prec G$ be strict divisions. If the subgraphs $H_1\Phi_1$ and $H_2\Phi_2$ in G are not* *disjoint, then there exists a connected transitive graph K such that H_1 and H_2 divide K and K strictly divides G.*

Proof: The proof is identical to that of Lemma 2.1 except that the graph K constructed from $H_1 \vee H_2$ requires a few more objects and arrows. Let a_1, a_2, b_1 and b_2 be defined as before, and let K be the transitive closure of the graph $H_1 \vee H_2$ with seven objects and eight arrows added. In the concrete situation where a_1 is a predecessor of b_1 and a_2 is a successor of b_2, the added objects $c_1, d_1, e_1, f, e_2, d_2$, and c_2 should have arrows added from c_1 to a_1, c_1 to d_1, e_1 to d_1, e_1 to f, f to e_2, d_2 to e_2, d_2 to c_2, and a_2 to c_2. Clearly $H_1 \vee H_2 \precsim K \precsim G$ since there is a faithful morphism $\Phi : K \to G$ which agrees with Φ_1 on the objects of H_1 and with Φ_2 on the objects of H_2 and which has $b_1\Phi = c_1\Phi = e_1\Phi$, $b_2\Phi = c_2\Phi = e_2\Phi$, and $a_1\Phi = d_1\Phi = f\Phi = d_2\Phi = a_2\Phi$. Notice that the objects c_i, d_i, and e_i for $i = 1, 2$ are either source or sink objects whose removal would disconnect the constructed graph and that they retain this property in the transitive closure K.

If K were equivalent to G, then by the Labeling Lemma there would exist a Φ-readable copy G' of G in K. Since G is connected, and since c_i and e_i are sent to the object $b_i\Phi$ in G, it follows that G' contains either c_1 or e_1 (but not both) and either c_2 or e_2 (but not both). Thus G' is contained in the full subgraph on $\mathrm{Obj}(H_1) \cup \{c_1, d_1\}$, the full subgraph on $\mathrm{Obj}(H_2) \cup \{c_2, d_2\}$, or the full subgraph on $\{d_1, e_1, f, e_2, d_2\}$. Since the first two cases are easily seen to be equivalent to H_1 and H_2, respectively, contradicting the assumption that Φ_1 and Φ_2 are strict divisions, the only possibility is that G' and thus G contains at most five objects and five edges. The lemma is easily seen to be true for the severely restricted possibilities which remain. □

Lemma 3.2 *If G is a connected and locally trivial transitive graph, then the family of connected transitive graphs strictly below G is upward-directed in the sense that, if H_1 and H_2 are connected and $\Phi_1 : H_1 \prec G$ and $\Phi_2 : H_2 \prec G$ are strict divisions, then there is a connected transitive graph K such that H_1 and H_2 divide K and K strictly divides G.*

Proof: The proof is identical to that of Lemma 2.2 once it is noted that the argument which shows that H_1' strictly divides G, also shows that the transitive closure of H_1' strictly divides G. Lemma 3.1 applied to the transitive closure of H_1' and H_2 completes the proof. □

Lemma 3.3 *If G is connected and $H_i \prec G$ for $i = 1, 2$ are gaps in* **FTG**, *then H_1 and H_2 are equivalent.*

Proof: The proof is identical to that of Lemma 2.3. □

3.2 Skeletons of Transitive Graphs

The skeleton of a finite transitive graph will be defined slightly differently from the skeleton of a directed graph in order to take advantage of the similarities between the two situations. A hom-set $G(a,b)$ will be called composite if there exist an object c and arrows $a \to c \to b$. If no such object c exists, then $G(a,b)$ is called skeletal. A composite or skeletal arrow is an arrow in a composite or skeletal hom-set. The **FTG**-skeleton of a transitive graph is the subgraph S which consists of all of the objects of G and exactly one arrow from each skeletal hom-set. Since the skeleton of a directed graph in the sense of Section 2 will no longer be needed, we may safely refer to the **FTG**-skeleton of a transitive graph G as simply the skeleton of G. Notice that the skeleton of a transitive graph is not itself a transitive graph, merely a directed one.

In addition to reformulating the notion of the skeleton of a graph, it is also necessary to reformulate the notions of a Φ-saturated graph, and a Φ-critical hom-set with respect to a faithful morphism Φ. Let $\Phi : H \to G$ be a faithful morphism between transitive graphs G and H. The graph H is called Φ-saturated if for every hom-set $H(a,b)$ of H which is not Φ-full, H strictly divides the transitive closure of $H + (a \to b)$.

The new definition of a Φ-critical hom-set requires an additional concept. A hom-set $G(c,d)$ is said to factor through $G(a,b)$ if there are arrows $c \to a$ and $b \to d$. This is equivalent, in a locally trivial transitive graph to the existence of a sequence $c = e_m \to e_{n-1} \ldots e_0 = a \to b = f_0 \to f_2 \to \ldots f_n = d$ where $G(e_i, e_{i-1})$ and $G(f_{j-1}, f_j)$ are skeletal hom-sets for all $i = 1, \ldots m-1$ and $j = 1, \ldots n-1$. This is true since if repeated factorization does not eventually stop at skeletal hom-sets, then an object will occur twice in the sequence, thereby, by transitivity producing a local arrow which contradicts our initial assumption. A hom-set $G(a,b)$ of a transitive graph G will be called Φ-critical if $H \prec G$ is a gap in **FTG**, $\Phi : H \to G$ is a faithful morphism, H is Φ-saturated, $G(a,b)$ has a pre-image under Φ which is not Φ-full, $S(a,b)$ splits the skeleton of G, and every composite hom-set $G(c,d)$ which factors through $G(a,b)$ contains exactly one arrow. A transitive graph G will be called trivial if G is locally trivial and all composite hom-sets have exactly one arrow. It should be noted that if all of the skeletal hom-sets in G are Φ-critical for some Φ, then G must be a trivial transitive graph.

Lemma 3.4 *If $\Phi : H \to G$ is a faithful morphism between transitive graphs and G is locally trivial, then there is a transitive graph H' which contains H as a subgraph, and there is an extension of Φ to a faithful morphism $\Phi' : H' \to G$, such that H' has the same objects as H, H' is equivalent to H, and H' is Φ'-saturated.*

Proof: If H is not Φ-saturated, then by definition there is a hom-set $H(a,b)$ which is not Φ-full and H is equivalent to the transitive closure of $H + (a \to b)$. In this case extend the faithful morphism Φ to the new arrows in the obvious way and repeat this procedure as many times as necessary. At some point the process must stop since the total number of arrows in any extension H' of H which divides G is bounded above by the number of hom-sets in H times the size of the largest hom-set in G. When it does stop, the graph and the faithful morphism under consideration satisfy the conclusion of the lemma. $\square$

Lemma 3.5 (Splitting Lemma) *Let G be a connected and reduced transitive graph, let $H \prec G$ be a gap in* **FTG***, and let $\Phi : H \to G$ be a faithful morphism with a surjective object function. If H is Φ-saturated and $H(a, b)$ is a hom-set which is not Φ-full, then its image under Φ, $G(a\Phi, b\Phi)$, is Φ-critical.*

Proof: We will first show that whenever a hom-set exists which is not Φ-full, there also exists a possibly different hom-set which is not Φ-full for which the addition of an arrow to this hom-set results in a transitive graph. Let $H(a, b)$ be a hom-set of H which is not Φ-full. If $H(a, b)$ is nonempty then this is immediate, since another arrow can be added to $H(a, b)$ without altering the fact that H is transitive. So assume $H(a, b)$ is empty. Next, define the influence of $H(a, b)$ to be the number of arrows which must be added to $H + (a \to b)$ to form its transitive closure. We will show that there are objects c and d for which $H(c, d)$ is not Φ-full and the influence of $H(c, d)$ is 0.

Assume that the influence of $H(a, b)$ is non-zero, and let $c \to d$ be one of the arrows which must be added to $H + (a \to b)$ to form its transitive closure. It is easy to see that the hom-set $H(c, d)$ is empty, that it factors through $H(a, b)$, and $H(c, d)$ is not Φ-full since $G(c\Phi, d\Phi)$ factors through $G(a\Phi, b\Phi)$ and is thus nonempty. In addition, since G, and thus H, are locally trivial, it follows that $H(a, b)$ does not factor through $H(c, d)$. Combining this with the fact that all of the hom-sets which factor through $H(c, d)$ also factor through $H(a, b)$, it follows that the influence of $H(c, d)$ is strictly smaller than that of $H(a, b)$. Continuing in this way we quickly find a hom-set satisfying the necessary conditions.

The rest of the proof follows the outline of the proof of Lemma 2.5. By the above argument, the hom-set $H(a, b)$ chosen at the beginning can now be chosen so that $H + (a \to b)$ is itself transitive. Another change is that the graph K which is constructed from $H \times \boldsymbol{n}^\bullet$ should be replaced by its transitive closure $\overline{K}$. The impossibility of an equivalence between $\overline{K}$ and H follows as before, so suppose that $\overline{K}$ is equivalent to G. As before, Lemma 1.3 implies that there is a Φ'-readable copy G' of G in $\overline{K}$. In this context this implies that G' contains exactly one object with first coordinate a and one object with first coordinate b. Next, notice that since composite hom-sets in $\overline{K}$ are sent under Φ' to composite hom-sets in G it follows that the inverse image under Φ' of a skeletal hom-set in G consists solely of skeletal hom-sets in $\overline{K}$. Thus the skeleton of G', which is isomorphic to the skeleton of G, is contained in the skeleton of $\overline{K}$. In particular, the only global skeletal hom-sets in K are between objects with first coordinates a and b.

Thus, if G' contains objects (a, i) and (b, i) for some i, then it cannot contain any of the objects (a, j) or (b, j) with $j \neq i$, and in particular, it cannot contain any of the global skeletal hom-sets of $\overline{K}$. Since the skeleton of G, like G itself, is connected, it follows that G divides the i-th level of K. But this level is isomorphic to H and we have reached a contradiction. Thus G' must contain (a, i) and (b, j) with $i \neq j$. Since G' cannot contain any of the other objects with first coordinate a or b, G' contains at most one global skeletal hom-set, namely, the hom-set $K((a, i), (b, j))$. If i is greater than j, then this hom-set is empty contradicting the connected nature of the skeleton of G and G'. On the other hand, if i is less than j then, since G' is isomorphic to G, it is clear that $G'((a, i), (b, j))$ splits the skeleton of G' and all of the composite hom-sets which factor through $G'((a, i), (b, j))$ contain exactly one arrow. Thus $G(a, b)$ splits the skeleton of G, and all of the composite hom-sets which factor through $G(a, b)$ contain exactly one arrow. In other words, $G(a, b)$ is Φ-critical. $\square$

Lemma 3.6 (Continued-splitting) *Let G be a connected and reduced transitive graph, let $H \prec G$ be a gap in* **FTG**, *let $\Phi : H \to G$ be a faithful morphism with a surjective object function, and let H be Φ-saturated. Suppose further that $G(a, b)$ is Φ-critical, and that there is object c in G such that one of the hom-sets $G(a, c)$, $G(c, a)$, $G(b, c)$, or $G(c, b)$ is nonempty. If Σ represents this nonempty hom-set, then there is a graph K which is equivalent to H, contains H as a full subgraph, and there is a faithful morphism $\Psi : K \to G$, such that K is Ψ-saturated and the hom-set Σ in K is Ψ-critical.*

Proof: The proof is identical to that of Lemma 2.6. It should be noted, however, that one of the consequences of Lemma 3.5 is that all hom-sets in H which are not Φ-full are skeletal hom-sets in G, and thus also skeletal hom-sets in H. In particular the graph H' constructed in the proof of Lemma 2.6 is already transitive. The proof can therefore be repeated without modification. $\square$

Lemma 3.7 *If G is a reduced and connected transitive graph and $H \prec G$ is a gap in* **FTG**, *then G is trivial and the skeleton of G is a tree.*

Proof: The proof is identical to that of Lemma 2.7 once it is observed that if all of the skeletal hom-sets in G are Φ-critical for some Φ, then G must be a trivial transitive graph. $\square$

3.3 Gaps between Connected Transitive Graphs

We shall now prove a converse of Lemma 3.7. Let G be a reduced, trivial, and connected transitive graph whose skeleton S is a tree. The construction given in Section 2 has been worded so that it applies equally well to the present situation. Notice in particular that a full link is added to all of the other possible links between a copy of a and a copy of any of the objects in the legacy of a. Notice also that since all of the non-full links correspond to skeletal hom-sets in G, $H(a_i, b_j)$ is nonempty whenever $G(a, b)$ is a composite hom-set. From this fact it follows easily that H_r is a transitive graph and thus so are all of its full subgraphs such as H_a and H. Once these observations have been made the rest of the results follow quickly. The complete construction has been carried out for $G = \vec{\mathbf{3}}$. See Figure 8.

Lemma 3.8 *Let G be an arbitrary reduced, connected, locally trivial, and nonempty transitive graph whose skeleton S is a tree. If H_r, and H are the transitive graphs constructed above, then H_r and G are equivalent, and H strictly divides G.*

Proof: The proof is identical to that of Lemma 2.8. □

Lemma 3.9 *Let G be an arbitrary reduced, connected, locally trivial, and nonempty transitive graph whose skeleton S is a tree and let H be the transitive graph constructed above. If K is any transitive graph such that K strictly divides G then K divides H. This immediately implies that $H \prec G$ is a gap in* **FTG**.

Proof: The proof is identical to that of Lemma 2.9. □

Theorem 3.10 *If G is a nonempty, connected, and reduced transitive graph, then there is a gap $H \prec G$ in* **FTG** *if and only if G is trivial and the skeleton of G is a tree. Moreover, when such a gap exists, the transitive graph H is unique up to equivalence.*

Proof: The theorem is a combination of Lemmas 3.3, 3.7, and 3.9. □

The properties of the set of gaps in **FTG** are similar to those in **FDG**. For example, it follows easily from the above construction that $\vec{\mathbf{1}} \prec \vec{\mathbf{2}} \prec \vec{\mathbf{3}} \prec \cdots$ are gaps in **FTG**. This was also proved directed by Rhodes in [6]. Since it is also true that $\emptyset \prec \bullet \prec E_1 \prec E_2 \prec \cdots$ are gaps in **FTG**, we have an easy example of a connected H which is at the bottom of two different gaps $H \prec G$ with G connected. Namely, let H be $E_1 = \vec{\mathbf{1}}$ and let G be either E_2 or $\vec{\mathbf{2}}$. A less immediate example, one where G_1 and G_2 are dual to each other and H is self-dual, is shown in Figure 9.

3.4 Gaps between Disconnected Transitive Graphs

As in the previous section we conclude our discussion of gaps in the category of finite transitive graphs by extending the above results on connected gaps to the disconnected ones. As before, the transition is fairly quick and immediate.

Lemma 3.11 *Let G and H be transitive graphs. If H strictly divides G, but H is not close to G, then $H \prec G$ is not a gap in* **FTG**.

Proof: The proof is identical to that of Lemma 2.11. □

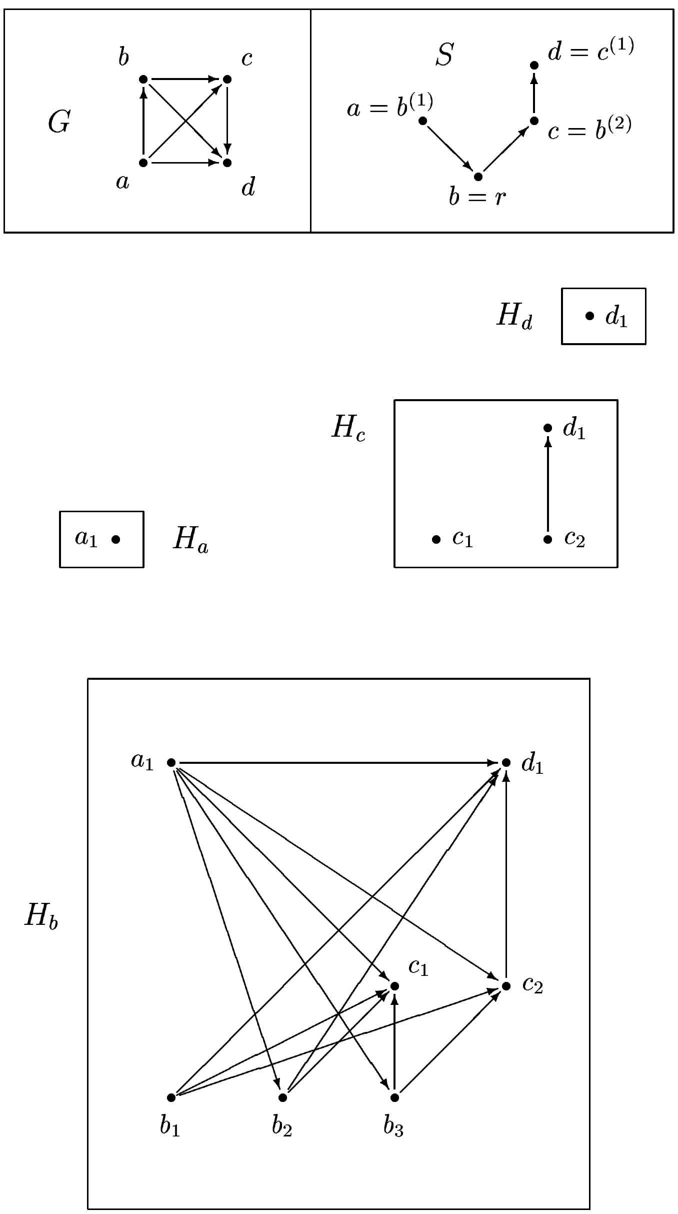

Figure 7: Sample construction of H_r. $G \cong \vec{\mathbf{3}}$ is shown in the upper left box. Object b is chosen as the root in S (upper right box). H will be obtained from H_b by deleting b_3 and all associated arrows. See Figure 8.

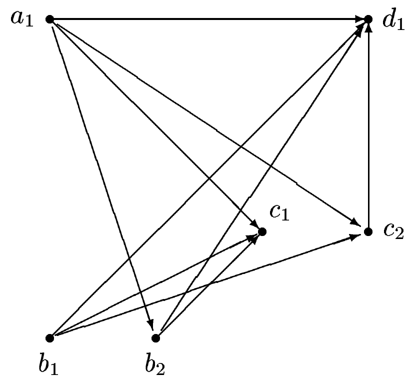

Figure 8: H constructed for the example of Figure 7, where G is isomorphic to $\vec{\mathbf{3}}$. This H is not reduced; it retracts to its full subgraph on b_1, c_2, and d_1.

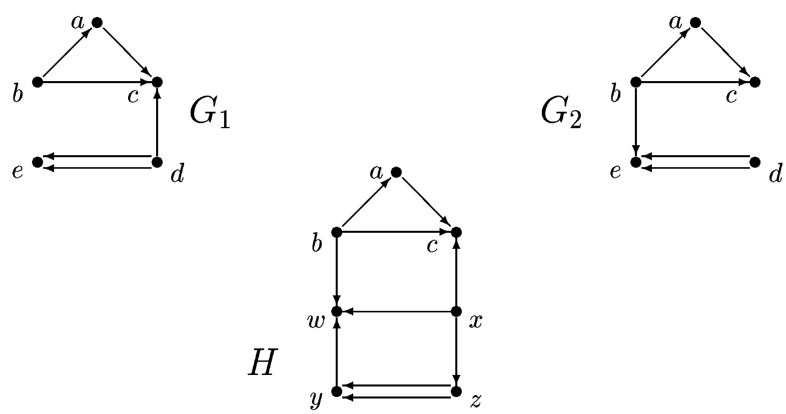

Figure 9: A connected H at the bottom of distinct gaps in **FTG**

Lemma 3.12 *Given transitive graphs $H \precsim K \precsim G$ with G reduced and H close to G, then there exists transitive graphs H', K', G', and L such that G' is connected, $H' \precsim K' \precsim G'$, and $H \sim H' \vee L$, $K \sim K' \vee L$, and $G \sim G' \vee L$.*

Proof: The proof is identical to that of Lemma 3.12. □

Theorem 3.13 *If G is a reduced transitive graph then there exists a transitive graph H for which $H \prec G$ is a gap in* **FTG** *if and only if G contains a connected component G' which is trivial and whose skeleton is a tree.*

Proof: As in the proof of Theorem 2.13, the sufficiency of the stated conditions is immediate. Thus it only remains to show that these conditions are necessary. Let $H \prec G$ be a gap. Since by Lemma 3.11 we may assume that H is close to G, let G' be the unique connected component of G which does not divide H and let L be the union of the other components. As above let H be written in the form $H' \vee L$ where H' does not divide L. All of the earlier arguments in this section can now be repeated with the graph L attached and with a few minor modifications to complete the proof. □

Exactly as in the case of finite directed graphs, the proof could actually be extended to give a one-to-one correspondence between the connected components of G which are trivial and whose skeletons are trees and those transitive graphs H for which $H \precsim G$ is a gap in **FTG**.

4 History and an Open Problem

The ordering under investigation in this article was originally defined on the set of finite categories by Bret Tilson in [7]. In [5], Rhodes first posed the question of a classification of the gaps in the Tilson ordering for finite undirected graphs, finite directed graphs, finite transitive graphs, and finite categories. In [6] he went on to show that when gaps exist in any of these categories, the upper part of the gap must be locally trivial, and he conjectured the following.

Conjecture 1 (Rhodes) *The divisions $\emptyset \prec \bullet \prec \boldsymbol{E_2}^\bullet$ are the only connected gaps in* **FCat**.

The next step forward came in [4] in which one of us (McCammond) announced new constructions which were thought to be sufficient to classify all of the gaps between finite categories. Although this has turned out not to be true, the announced constructions did form the basis for articles ([2] and the present article) which have succeeded in classifying the gaps between finite undirected, finite directed, and finite transitive graphs. In [1], significant partial results have been obtained for gaps between finite categories. In particular, if G is a reduced and connected finite category which does not divide $\boldsymbol{E_2}^\bullet$ and $H \prec G$ a gap in **FCat**, then it has been shown that the underlying transitive graph of G with the identity arrows removed must be locally trivial and trivial, G cannot contain $E_{2,2}$ as a subgraph, the hom-sets of G must factor uniquely into skeletal hom-sets, G must contain strictly more than two objects, and at least one of the objects must have a proper predecessor and a proper successor. In spite of this progress, Rhodes' conjecture remains open.

References

[1] R. Baldinger. *Gaps in Finite Directed Structures.* PhD thesis, U.C. Berkeley, 1994.

[2] R. Baldinger and J. McCammond. Gaps between connected finite graphs. *International Journal of Algebra and Computation*, 2:19–31, 1992.

[3] S. MacLane. *Categories for the Working Mathematician.* Springer-Verlag, New York, 1971.

[4] J. McCammond. Gaps in categories. In J. Rhodes, editor, *Monoids and Semigroups with Applications*, pages 83–85. World Scientific, Singapore, 1991.

[5] J. Rhodes. New techniques in global semigroup theory. In S. Goberstein and P. Higgins, editors, *Semigroups and Their Applications*, pages 169–181. Reidel, Dordrecht, 1987.

[6] J. Rhodes. On the Cantor-Dedekind property of the Tilson order on categories and graphs. *Journal of Pure and Applied Algebra*, 48:55–82, 1987.

[7] B. Tilson. Categories as algebra: An essential ingredient in the theory of monoids. *Journal of Pure and Applied Algebra*, 48:83–198, 1987.