Down-coloring acyclic digraphs

Geir Agnarsson * Ágúst S. Egilsson † Magnús M. Halldórsson ‡

Abstract

We consider vertex coloring of an acyclic digraph $\vec{G}$ in such a way that two vertices which have a common ancestor in $\vec{G}$ receive distinct colors. Such colorings arise in a natural way when bounding space for various genetic data for efficient analysis. We discuss the corresponding chromatic number and derive an upper bound as a function of the maximum number of descendants of a given vertex and the degeneracy of the corresponding hypergraph, which is obtained from the original digraph.

2000 MSC: 05C15, 05C20, 05C65, 05C85, 05D90, 06A05.

Keywords: Genetic databases, compact matrix representation, digraph, transitive closure, ancestor, down-set, vertex coloring, hypergraph.

1 Introduction

The purpose of this article is to discuss a special kind of vertex coloring for acyclic digraphs, where vertices with a common ancestor must receive distinct colors. We discuss some properties of such colorings, similarity and differences with ordinary graph coloring, and derive an upper bound which, in addition, yields an efficient coloring procedure. We will give some different representations of our acyclic digraph, some equivalent and other more relaxed representations, which, from our vertex coloring point of view, will suffice to consider.

Digraphs representing various biological phenomena and knowledge are ubiquitous in the life sciences and in drug discovery research, e.g. the gene ontology digraph maintained by the Gene Ontology Consortium [5]. An overview of several projects relating to indexing of semistructured data (i.e. acyclic digraphs) can be found in [1]. In these biological digraphs it is important to be able to access the ancestors of nodes in a fast and efficient manner.

Consider the problem of finding a representation an acyclic digraph in a database to allow for fast access to the set of ancestors of a given node. The ancestors of a node are its

*Department of Mathematical Sciences, George Mason University, MS 3F2, 4400 University Drive, Fairfax, VA 22030, geir@math.gmu.edu.

†Department of Mathematics, University of California, Berkeley, CA 94720-3840, egilsson@math.berkeley.edu.

‡Department of Computer Science, University of Iceland, IS-107 Rvk, Iceland, mmh@hi.is.

in-neighbors in the transitive closure of the digraph. If the digraph is sparse and shallow, the transitive closure is also sparse. Thus, an adjacency matrix representation would be neither efficient nor fast. On the other hand, a matrix has the advantage of corresponding nicely to the relational representation used in modern databases. In a database relation corresponding to an adjacency matrix, the non-empty elements in each column correspond to the in-neighbors of the node indexing that column. This vertex set can then be combined by relational joins with other tables that are also indexed by vertices, giving an effective language of querying based on graphical properties. Thus, it would be preferable to find a matrix representation that consists of relatively small rows.

One compact matrix representation would be to store the adjacency lists in compacted array form, where a list with k elements is stored in the first k array elements. In this case, however, there is no easy way of accessing all the edges entering a given vertex. While this could be alleviated by storing the inverted adjacency matrix, note that in the context of database access, rows are conceptually different from columns. Instead, we seek a compacted representation where all in-edges of a given node are stored in the same column. We say that a many-to-one mapping of vertices to columns that preserves adjacency lists, has the *adjacency-consistent property.* By recording the mapping of each node to the column storing its in-neighbors, one obtains the same desirable properties of adjacency matrices in the context of a relational database. If the graph is sparse, the possibilities of storage reduction are significant. The actual improvement is related to the number of colors needed in a certain coloring of the digraph, which we now describe.

A proper *down-coloring* of a digraph is a vertex coloring where vertices with a common ancestor receive different colors. The *down-chromatic number* of a digraph is the minimum number of colors in a down-coloring of the digraph. In a compacted matrix representation of the transitive closure of a digraph, we assign multiple vertices to the same column, but in such a way that their in-adjacency lists must be disjoint. Two vertices have disjoint sets of in-neighbors in the transitive closure if, and only if, they have no common ancestor. Therefore, a down-coloring of a digraph corresponds to a valid compacted representation of its transitive closure, and the down-chromatic number is the minimum number of columns needed in such a representation.

Example: Consider the digraph $\vec{G}$, on $n = 6$ vertices representing genes, where a directed edge from one vertex to a second one indicates that the first gene is an ancestor of the second gene.

$$
\begin{aligned}
V(\vec{G}) &= \{g_1, g_2, g_3, g_4, g_5, g_6\},\\
E(\vec{G}) &= \{(g_1, g_4), (g_1, g_5), (g_2, g_4), (g_2, g_6), (g_3, g_5), (g_3, g_6)\}.
\end{aligned}
$$

In the adjacency matrix representation of this 6 node digraph, we assign a column to each vertex g_i. As we see in the left diagram of Table 1, most of the entries of this 6×6 matrix are empty. By reducing the number of columns in such a way that the adjacency-consistent property still holds, we obtain a smaller and more compact 6×3 matrix representation as seen on the right diagram of Table 1. There, the i-th row still contains the descendants of

	g_1	g_2	g_3	g_4	g_5	g_6
g_1	1	0	0	1	1	0
g_2	0	1	0	1	0	1
g_3	0	0	1	0	1	1
g_4	0	0	0	1	0	0
g_5	0	0	0	0	1	0
g_6	0	0	0	0	0	1

$\longrightarrow$

	1	2	3
g_1	g_1	g_4	g_5
g_2	g_6	g_4	g_2
g_3	g_6	g_3	g_5
g_4	–	g_4	–
g_5	–	–	g_5
g_6	g_6	–	–

Table 1: An example for $n = 6$.

g_i and the ancestors of g_i are those g_j's whose rows g_i appears in. Note that, (i) each g_i appears in exactly one column and, (ii) two genes appear in the same column only if their sets of ancestors are disjoint. Here we can view the column numbers 1, 2, and 3 as distinct colors assigned to each vertex. Note further that in this example the transitive closure of $\vec{G}$ is simply $\vec{G}$ itself.

An explicit example on how such a coloring can speed up queries in the gene ontology digraph can be found in the appendix of [4].

1.1 Related Work

This article can be viewed as a continuation of [2] where some special classes of digraphs are studied, in particular those of height two in which every vertex has an in-degree of two. For a brief introduction and additional references to the ones mention here, we refer to [2].

Directed acyclic graphs are frequently called DAG's by computer scientists, as is the case in [8, p. 194], but we will here call them acyclic digraphs.

A straight forward condition of a vertex coloring of a digraph $\vec{G}$ is to insist that two vertices u and v recieve distinct color if there is a directed edge from u to v in $\vec{G}$. Such a coloring is, of course, the same as coloring the vertices of the underlying graph G of $\vec{G}$ (by forgetting the orientation of the directed edges) in the usual sence. Only few vertex coloring results however rely on the structure that the directions of the edges in a digraph provides. We describe the two most common ones that appear in the mathematics and computer science literature:

The *arc-chromatic number* of a digraph $\vec{G}$, as investigated in [6], is defined as follows: First form the *arc digraph* or the *line digraph* $\vec{L}(\vec{G})$ of $\vec{G}$ by letting the vertices of $\vec{L}(\vec{G})$ consist of the directed edges (u, v) of $\vec{G}$, and a directed edge from (u, v) to (u', v') is present in $\vec{L}(\vec{G})$ if, and only if, $v = u'$. The arc-chromatic number of $\vec{G}$ is then defined to be the chromatic number of $L(\vec{G})$, the underlying graph of the arc-digraph $\vec{L}(\vec{G})$ of G. Note that although the chromatic number of $L(\vec{G})$ does not depend on the direction of the edges of $\vec{L}(\vec{G})$, the definition of the arc-digraph $\vec{L}(\vec{G})$ itself does rely on the direction of the edges of $\vec{G}$.

Another vertex coloring of digraphs that relies on the direction of the egdes is the *dichromatic number* of a digraph $\vec{G}$, as studied in [7] and [9], which is defined as the minimum number of colors needed to vertex color $\vec{G}$ in such a way that no monochromatic directed cycle is created.

In the following section we state the basic definitions and observations needed for the rest of the paper.

2 Definitions and observations

We attempt to be consistent with standard graph theory notation in [12], and the notation in [10] when applicable.

For a natural number $n \in \mathbb{N}$ we let $[n] = \{1, \ldots, n\}$. A *simple digraph* is a finite simple directed graph $\vec{G} = (V, E)$, where $V = V(\vec{G})$ is a finite set of vertices and $E = E(\vec{G}) \subseteq V \times V$ is a set of directed edges. The digraph $\vec{G}$ is said to be *acyclic* if $\vec{G}$ has no directed cycles. Henceforth $\vec{G}$ will denote an acyclic digraph in this section.

The binary relation $\leq$ on $V(\vec{G})$ defined by

$$u \leq v \Leftrightarrow u = v, \text{ or there is a directed path from } v \text{ to } u \text{ in } \vec{G}, \tag{1}$$

is reflexive, antisymmetric and transitive and therefore a partial order on $V(\vec{G})$. Hence, whenever we talk about $\vec{G}$ as a *poset*, the partial order will be the one defined by (1). Note that the acyclicity of $\vec{G}$ is essential in order to be able to view $\vec{G}$ as a poset. The *transitive closure* of $\vec{G}$ is the poset $\vec{G}$ viewed as a digraph, that is the digraph $\vec{G}^*$ on $V(\vec{G})$ where $(v, u) \in E(\vec{G}^*)$ iff $u < v$. By the *height* of $\vec{G}$ as a poset, we mean the number of vertices in the longest directed path in $\vec{G}$. We denote by $\max\{\vec{G}\}$ the set of maximal vertices of $\vec{G}$ with respect to the partial order $\leq$.

For vertices $u, v \in V(\vec{G})$ with $u \leq v$, we say that u is a *descendant* of v, and v is an *ancestor* of u. The *down-set* $D[u]$ of a vertex $u \in V(\vec{G})$ is the set of descendants of u in $\vec{G}$, that is, $D[u] = \{x \in V(\vec{G}) : x \leq u\}$.

Definition 2.1 *A* down-coloring *of $\vec{G}$ is a map $c : V(\vec{G}) \to [k]$ satisfying*

$$u, v \in D[w] \textit{ for some } w \in V(\vec{G}) \ \Rightarrow c(u) \neq c(v)$$

for every $u, v \in V(\vec{G})$. The down-chromatic number *of $\vec{G}$, denoted by $\chi_d(\vec{G})$, is the least k for which $\vec{G}$ has a proper down-coloring $c : V(\vec{G}) \to [k]$.*

Just as in an undirected graph G, the vertices in a clique must all receive distinct colors in a proper vertex coloring of G. Therefore $\omega(G) \leq \chi(G) \leq |V(G)|$ where $\omega(G)$ denotes the clique number of G. In addition $\chi(G)$ can be larger than the clique number $\omega(G)$. Similarly for our acyclic digraph $\vec{G}$ we have $D(\vec{G}) \leq \chi_d(\vec{G}) \leq |V(\vec{G})|$, where we define

$$D(\vec{G}) = \max_{u \in V(\vec{G})} \{|D[u]|\}.$$

We will see below that the same holds for $\vec{G}$, that $\chi_d(\vec{G})$ can be much larger than $D(\vec{G})$.

When considering down-colorings, it can be useful to map the problem to one on undirected graphs. Given an acyclic digraph $\vec{G}$, the corresponding simple undirected *down-graph* G' has the same set of vertices, with each pair of vertices connected that are contained in the same principal down-set:

$$\begin{aligned} V(G') &= V(\vec{G}), \\ E(G') &= \{\{u,v\} : u, v \in D[w] \text{ for some } w \in V(\vec{G})\}. \end{aligned}$$

In this way we have transformed the problem of down-coloring the digraph $\vec{G}$ to the problem of vertex coloring the simple undirected graph G' in the usual sense, and we have $\chi_d(\vec{G}) = \chi(G')$. Hence, from the point of down-colorings, both $\vec{G}$ and G' are equivalent, which is something we will discuss in the next section to come. However, some structure is lost. The fact that two vertices u and v are connected in G' could mean one of three possibilities, (i) $u < v$, (ii) $u > v$, or (iii) u and v are incomparable, but there is a vertex w with $u < w$ and $v < w$.

Although a down-set in $\vec{G}$ will become a clique in G', the converse is not true, as stated in Observation 2.2 below.

2.1 Relating the down-chromatic number to the size of down-sets

We now consider a concrete example and some observations drawn from it. A special case of this following example can be found in [2].

EXAMPLE: Let $k \geq 2$ and $m \geq 1$ be natural numbers. Let $A_1, \dots, A_k$ be disjoint sets, each A_i containing exactly m vertices. For each of the $\binom{k}{2}$ pairs $\{i,j\} \subseteq [k]$ define an additional vertex w_{ij}. Let $\vec{G}(k,m)$ be a digraph with vertex set and edge set given by

$$\begin{aligned} V(\vec{G}(k,m)) &= \left(\bigcup_{i \in [k]} A_i \right) \cup \{w_{ij} : \{i,j\} \subseteq [k]\}, \\ E(\vec{G}(k,m)) &= \bigcup_{\{i,j\} \subseteq [k]} \{(w_{ij}, u) : u \in A_i \cup A_j\}. \end{aligned}$$

Clearly $\vec{G}(k,m)$ is a simple acyclic digraph on $n = km + \binom{k}{2}$ vertices and with $\binom{k}{2} \cdot 2m = k(k-1)m$ directed edges. Each closed principal down-set is of the form $D[w_{ij}]$ for some $\{i,j\} \subseteq [k]$ and hence $D(\vec{G}(k,m)) = 2m+1$. Note that in any proper down-coloring of $\vec{G}(k,m)$, every two vertices in $\bigcup_{i \in [k]} A_i$ are both contained in $D[w_{ij}]$ for some $\{i,j\} \subseteq [k]$, and hence $\bigcup_{i \in [k]} A_i$ forms a clique in $G'(k,m)$. From this we see that $\omega(G'(k,m)) = km$. In particular we have that

$$\chi_d(\vec{G}(k,m)) = \chi(G'(k,m)) \geq \omega(G'(k,m)) = km.$$

In fact, any coloring of $\bigcup_{i \in [k]} A_i$ with km colors can be extended in a greedy fashion to a proper down-coloring of $\vec{G}(k,m)$ with at most km colors. Therefore equality holds through the above display. In particular we have $D(\vec{G}(k,1)) = 3$ and $\omega(G'(k,1)) = k$, from which we deduce the following observation.

Observation 2.2 *There is no function $f : \mathbb{N} \to \mathbb{N}$ with $\omega(G') \le f(D(\vec{G}))$ for every acyclic digraph $\vec{G}$. In particular, there is no function f with $\chi_d(\vec{G}) \le f(D(\vec{G}))$ for all acyclic digraphs $\vec{G}$.*

Let $\alpha \in \mathbb{N}$ be a fixed and "large" natural number. Denoting by n the number of vertices of $\vec{G}(k, \alpha k)$, we clearly have $n = |V(\vec{G}(k, \alpha k))| = \alpha k^2 + \binom{k}{2} \sim k^2(\alpha + 1/2)$, where $f(k) \sim g(k)$ means $\lim_{k\to\infty} f(k)/g(k) = 1$. We now see that

$$\begin{aligned} D(\vec{G}(k,\alpha k)) &= 2\alpha k + 1 \sim \left(\frac{2\alpha}{\sqrt{\alpha+1/2}}\right)\sqrt{n}, \\ \chi_d(\vec{G}(k,\alpha k)) &= \alpha k^2 \sim \left(\frac{\alpha}{\alpha+1/2}\right) n. \end{aligned}$$

From this we have the following.

Observation 2.3 *For every $\epsilon > 0$, there is an $n \in \mathbb{N}$ for which there is an acyclic digraph $\vec{G}$ on n vertices with $D(\vec{G}) = \Theta(\sqrt{n})$ and $\chi_d(\vec{G}) \ge (1-\epsilon)n$.*

REMARK: The above Observation 2.3 simply states that $\chi_d(\vec{G})$ can be an arbitrarily large fraction of $|V(\vec{G})|$ without making $D(\vec{G})$ too large. Hence, both the upper and lower bound in the inequality $D(\vec{G}) \le \chi_d(\vec{G}) \le |V(\vec{G})|$ are tight in this sense.

3 Hypergraph representations

In this section we discuss alternative representations of our digraph $\vec{G}$, and define some parameters which we will use to bound the down-chromatic number $\chi_d(\vec{G})$.

We first consider the issue of the height of digraphs. We say that two digraphs on the same set of vertices are *equivalent* if every down-coloring of one is also a valid down-coloring of the other, that is, if they induce the same undirected down-graph. We show that for any acyclic digraph $\vec{G}$ there is an equivalent acyclic digraph $\vec{G}_2$ of height two with $\chi_d(\vec{G}) = \chi_d(\vec{G}_2)$. However, the degrees of vertices in $\vec{G}_2$ may necessarily be larger than in $\vec{G}$.

Proposition 3.1 *Any down-graph G' of an acyclic digraph $\vec{G}$ is also a down-graph of an acyclic digraph $\vec{G}_2$ of height two.*

Proof. The derived digraph $\vec{G}_2$ has the same vertex set as $\vec{G}$, while the edges all go from $\max\{\vec{G}\}$ to $V(\vec{G}) \setminus \max\{\vec{G}\}$, where $(u,v) \in E(\vec{G}_2)$ if, and only if, $v \in D(u)$. In this way we see that two vertices in $\vec{G}$ have a common ancestor if, and only if they have a common ancestor in $\vec{G}_2$. Hence, we have the proposition. $\square$

Therefore, when considering down-colorings of digraphs, we can by Proposition 3.1 assume them to be of height two.

There is a natural correspondence between acyclic digraphs and certain hypergraphs.

Definition 3.2 *For a digraph $\vec{G}$, the related* down-hypergraph $H_{\vec{G}}$ *of $\vec{G}$ is given by:*

$$\begin{aligned} V(H_{\vec{G}}) &= V(\vec{G}) \\ \mathcal{E}(H_{\vec{G}}) &= \{D[u] : u \in \max\{\vec{G}\}\}. \end{aligned}$$

Note that the down-graph G' is the *clique graph* of the down-hypergraph $H_{\vec{G}}$, that is, the undirected graph where every pair of vertices which are contained in a common hyperedge in $H_{\vec{G}}$ are connected by an edge in G'.

As we shall see, not every hypergraph is a down-hypergraph. There is a simple criterion for whether a hypergraph is a down-hypergraph or not. An edge in a hypergraph has a *unique-element* if it contains a vertex contained in no other edge. A hypergraph has the *unique-element property* if every edge has a unique element.

Observation 3.3 *A hypergraph is a down-hypergraph if, and only if, it has the unique-element property.*

Proof. A down-hypergraph is defined to contain the principal closed down-sets of a digraph $\vec{G}$ as edges. Each such edge contains a maximal element in $\vec{G}$, and this element is not contained in any other down-set. Hence, a down-hypergraph satisfies the unique-element property.

Suppose a hypergraph H satisfies the property. Then we can form a height-two acyclic digraph as follows: For each hyperedge, add a source vertex in the digraph corresponding to the representative unique element of the hyperedge. For the other hypervertices, add sinks to the digraph with edges from the sources to those sinks that correspond to vertices in the same hyperedge. □

Note that a hypergraph with the unique element property is necessarily *simple*, in the sense that each hyperedge is uniquely determined by the vertices it contains.

We see that we can convert a proper down-hypergraph to a corresponding acyclic digraph of height two, and vice versa, in polynomial time.

Definition 3.4 *A* strong coloring *of a hypergraph H, is a map $\Psi : V(H) \to [k]$ satisfying*

$$u, v \in e \textit{ for some } e \in \mathcal{E}(H) \Rightarrow \Psi(u) \neq \Psi(v).$$

The strong chromatic number $\chi_s(H)$ *is the least number k of colors for which H has a proper strong coloring $\Psi : V(H) \to [k]$.*

Note that a strong coloring of a down-hypergraph $H_{\vec{G}}$ is equivalent to a down-coloring of $\vec{G}$, and hence $\chi_s(H_{\vec{G}}) = \chi_d(\vec{G})$. Since we can convert to and from hypergraph and digraph representations, the two coloring problems are polynomial-time reducible to each other.

Strong colorings of hypergraphs have been studied, but not to the extent of various other types of colorings of hypergraphs. In [11] a nice survey of various aspects of hypergraph coloring theory is found, containing almost all fundamental results in the past three decades.

In the next section we will bound the down-chromatic number of our acyclic digraph $\vec{G}$, partly by another parameter of the corresponding down-hypergraph $H_{\vec{G}}$.

4 Upper bound in terms of degeneracy

As we saw in Observation 2.2, it is in general impossible to bound $\chi_d(\vec{G})$ from above solely in terms of $D(\vec{G})$, even if $\vec{G}$ is of height two. Therefore we need an additional parameter for that upper bound, but first we need to review some matters about a hypergraph $H = (V(H), \mathcal{E}(H))$.

Two vertices in $V(H)$ are *neighbors* in H if they are contained in the same edge in $\mathcal{E}(H)$. An edge in $\mathcal{E}(H)$ containing just one element is called *trivial*. The largest cardinality of a hyperedge of H will be denoted by $\sigma(H)$. The *degree* $d_H(u)$, or just $d(u)$, of a vertex $u \in V(H)$ is the number of non-trivial edges containing u. Note that $d(u)$ is generally much smaller than the number of neighbors of u. The minimum and maximum degree of H are given by

$$\begin{aligned} \delta(H) &= \min_{u \in V(H)} \{d_H(u)\}, \\ \Delta(H) &= \max_{u \in V(H)} \{d_H(u)\}. \end{aligned}$$

The subhypergraph $H[S]$ of H, induced by a set S of vertices, is given by

$$\begin{aligned} V(H[S]) &= S, \\ \mathcal{E}(H[S]) &= \{X \cap S : X \in \mathcal{E}(H) \text{ and } |X \cap S| \geq 2\}. \end{aligned}$$

Definition 4.1 *Let H be a simple hypergraph. The* degeneracy *or the* inductiveness *of H, denoted by* $\mathrm{ind}(H)$*, is given by*

$$\mathrm{ind}(H) = \max_{S \subseteq V(H)} \{\delta(H[S])\}.$$

If $k \geq \mathrm{ind}(H)$, then we say that H is k-degenerate *or* k-inductive.

Note that Definition 4.1 is a generalization of the degeneracy or the inductiveness of a usual undirected graph G, given by $\mathrm{ind}(G) = \max_{H \subseteq G} \{\delta(H)\}$, where H runs through all the induced subgraphs of G. Note that the degeneracy of a (hyper)graph is always greater than or equal to the degeneracy of any of its sub(hyper)graphs.

To illustrate, let us for a brief moment discuss the degeneracy of an important class of simple graphs, namely that of simple planar graphs. Every subgraph of a simple planar graph is again planar. Since every planar graph has a vertex of degree five or less, the degeneracy of every planar graph is at most five. This is the best possible for planar graphs, since the graph of the icosahedron is planar and 5-regular. That a planar graph has degeneracy of five, implies that it can be vertex colored in a simple greedy fashion with at most six colors. The degeneracy has also been used to bound the chromatic number of the square G^2 of a planar graph G, where G^2 is a graph obtained from G by connecting two vertices of G if, and only if, they are connected in G or they have a common neighbor in G, [3].

In general, the degeneracy of an undirected graph G yields an ordering $\{u_1, u_2, \ldots, u_n\}$ of $V(G)$, such that each vertex u_i has at most $\mathrm{ind}(G)$ neighbors among the previously listed vertices $u_1, \ldots, u_{i-1}$. Such an ordering provides a way to vertex color G with at most $\mathrm{ind}(G) + 1$ colors in an efficient greedy way, and hence we have in general that $\chi(G) \leq \mathrm{ind}(G) + 1$.

The degeneracy of a simple hypergraph is also connected to a greedy vertex coloring of it, but not in such a direct manner as for a regular undirected graph, since, as noted, the number of neighbors of a given vertex in a hypergraph is generally much larger than its degree.

Theorem 4.2 *If the simple undirected graph G is the clique graph of the simple hypergraph H then* $\mathrm{ind}(G) \leq \mathrm{ind}(H)(\sigma(H) - 1)$.

Proof. For each $S \subseteq V(G) = V(H)$, let $G[S]$ and $H[S]$ be the subgraph of G and the subhypergraph of H induced by S, respectively. Note that for each $u \in S$, each hyperedge in $H[S]$ which contains u, has at most $\sigma(H[S]) - 1 \leq \sigma(H) - 1$ other vertices in addition to u. By definition of $d_{H[S]}(u)$, we therefore have that $d_{G[S]}(u) \leq d_{H[S]}(u)(\sigma(H) - 1)$, and hence

$$\delta(G[S]) \leq \delta(H[S])(\sigma(H) - 1). \tag{2}$$

Taking the maximum of (2) among all $S \subseteq V(G)$ yields the theorem. □

Recall that the *intersection graph* of a collection $\{A_1, \ldots, A_n\}$ of sets, is the simple graph with vertices $\{u_1, \ldots, u_n\}$, where we connect u_i and u_j if, and only if, $A_i \cap A_j \neq \emptyset$.

Observation 4.3 *For a simple connected hypergraph H, then* $\mathrm{ind}(H) = 1$ *if, and only if, the intersection graph of its hyperedges $\mathcal{E}(H)$ is a tree.*

What Observation 4.3 implies, is that edges of H can be ordered as $\mathcal{E}(H) = \{e_1, \ldots, e_m\}$, such that each e_i intersects exactly one edge from the set $\{e_1, \ldots, e_{i-1}\}$. If now G is the clique graph of H, this implies that $\mathrm{ind}(G) = \sigma(H) - 1$.

Also note that if H has the unique element property and $\sigma(H) = 2$, then clearly the clique graph G is a tree, and hence $\mathrm{ind}(G) = 1 = \sigma(H) - 1$. We summarize in the following.

Observation 4.4 *Let H be a hypergraph that satisfies the unique element property. If either* $\mathrm{ind}(H) = 1$ *or* $\sigma(H) = 2$*, then the clique graph G of H satisfies* $\mathrm{ind}(G) = \sigma(H) - 1$.

For a hypergraph H with the unique element property, we can obtain some slight improvements in the general case as well.

Theorem 4.5 *Let H be a hypergraph with the unique element property. Assume further that* $\mathrm{ind}(H) > 1$ *and* $\sigma(H) > 2$*. Then the graph G of H satisfies*

$$\mathrm{ind}(G) \leq \mathrm{ind}(H)(\sigma(H) - 2).$$

Proof. Since H has the unique element property, then by Observation 3.3 there is an acyclic digraph $\vec{G}$ such that $H = H_{\vec{G}}$. Let H'' be the hypergraph induced by $V(H) \setminus \max\{\vec{G}\}$ and G'' be the corresponding clique graph of H''. Since each $u \in \max\{\vec{G}\}$ is simplicial in H and in G, their removal will not effect the degeneracy of the remaining vertices, so $\mathrm{ind}(H'') = \mathrm{ind}(H)$ and $\mathrm{ind}(G'') = \mathrm{ind}(G)$. Also note that $\sigma(G'') = \sigma(G) - 1$. By Theorem 4.2 we get $\mathrm{ind}(G) = \mathrm{ind}(G'') \leq \mathrm{ind}(H'')(\sigma(H'') - 1) = \mathrm{ind}(H)(\sigma(H) - 2)$, thereby completing the proof. □

Let $\vec{G}$ be an acyclic digraph. Since now $D(\vec{G}) = \sigma(H_{\vec{G}})$ and $\chi_d(\vec{G}) = \chi(G') \leq \mathrm{ind}(G')+1$, we obtain the following summarizing corollary.

Corollary 4.6 *If $\vec{G}$ is an acyclic digraph, then its down-chromatic number satisfies the following:*

1. *If* $\mathrm{ind}(H_{\vec{G}}) = 1$ *or* $D(\vec{G}) = 2$*, then* $\chi_d(\vec{G}) = D(\vec{G})$*.*
2. *If* $\mathrm{ind}(H_{\vec{G}}) > 1$ *and* $D(\vec{G}) > 2$*, then* $\chi_d(\vec{G}) \leq \mathrm{ind}(H_{\vec{G}})(D(\vec{G}) - 2) + 1$*.*

Moreover, in both cases the given upper bound of colors can be used do down-color $\vec{G}$ in an efficient greedy fashion.

Example: Let $k, m \in \mathbb{N}$, assuming them to be "large" numbers. Consider the graph $\vec{G}(k, m))$ from Section 2. Here we have that $H_{\vec{G}(k,m)}$ is a hypergraph with $\mathrm{ind}(H_{\vec{G}(k,m)}) = 2(k-2)$ and $\sigma(H_{\vec{G}(k,m)}) = 2m+1$. By Corollary 4.6 we have immediately that $\chi_d(\vec{G}(k, m)) \leq 2(k-2)(2m-1)+1 = \Theta(km)$, which agrees with the asymptotic value of the actual down-chromatic number km (also a $\Theta(km)$ function,) which we computed in Section 2. Hence, Corollary 4.6 is asymptotically tight.

Moreover, if we were to just color each vertex with its own unique color, we compare $\chi_d(\vec{G}(k, m)) \leq 2(k-2)(2m-1)+1$ with the actual number $km + \binom{k}{2}$ of vertices, and we see that for large k, this is a substantial reduction.

Remark: Although we have assumed our digraphs to be acyclic, we note that the definition of down-coloring can be easily extended to a regular cyclic digraph $\vec{G}$ by interpreting the notion of descendants of a vertex u to mean the set of nodes reachable from u. In fact, if $\vec{G}$ is an arbitrary digraph, then there is an equivalent acyclic digraph $\vec{G}'$, on the same set of vertices, with an identical down-graph: First form the *condensation* $\hat{G}$ of $\vec{G}$ by shrinking each strongly connected component of $\vec{G}$ to a single vertex. Then form $\vec{G}'$ by replacing each node of $\hat{G}$ which represents a strongly connected component of $\vec{G}$ on a set $X \subseteq V(\vec{G})$

of vertices, with an arbitrary vertex $u \in X$, and then add a directed edge from u to each $v \in X \setminus \{u\}$. This completes the construction.

Observe that each node $v \in X$ has exactly the same neighbors in the down-graph of $\vec{G}'$ as u, as it is a descendant of u and u alone. Further, if node v was in a different strong component of $\vec{G}$ than u but was reachable from u, then it will continue to be a descendant of u in $\vec{G}'$. Hence, the down-graphs of $\vec{G}$ and $\vec{G}'$ are identical.

Acknowledgments

The authors are grateful to John L. Pfaltz for his interest and encouragements to write this article.

References

[1] S. Abitebboul, P. Buneman and D. Suciu. Data on the web, from relations to semistructured data and XML. *Morgan Kaufmann Publishers*, (2000).

[2] G. Agnarsson and Á. Egilsson. On vertex coloring simple genetic digraphs. *Congressus Numerantium*, **161**, 117 – 127, (2004).

[3] G. Agnarsson and M. M. Halldórsson, Coloring powers of planar graphs, *SIAM Journal of Discrete Mathematics*, **16(4)**, 651 – 662, (2003).

[4] Á. Egilsson, H. Gudbjartsson. Indexing of tables referencing complex structures. preprint. Available at `http://arxiv.org/abs/cs.DB/0309011`, (2003).

[5] The Gene Ontology Consortium. Gene Ontology: Tool for the unification of biology. *Nature Genet*, **25**, 25 – 29, (2000).

[6] C. C. Harner and R. C. Entringer. Arc colorings of digraphs. *Journal of Combinatorial Theory, Series B*, **13**, 219 – 225, (1972).

[7] H. Jacob and H. Meyniel. Extensions of Turán's Brooks' theorems and new notions of stability and colorings in digraphs. *Combinatorial Mathematics, North-Holland Math. Stud.*, **75**, 365 – 370, (1983).

[8] C. A. Shaffer. A practical introduction to data structures and algorithm analysis *Java Edition, Prentice Hall*, (1998).

[9] X. Y. Su. Brooks' theorem on colorings of digraphs. *Fujian Shifan Daxue Xuebao Ziran Kexue Ban*, **3(1)**, 1 – 2, (1987).

[10] W. T. Trotter. Combinatorics and partially ordered sets, dimension theory. *Johns Hopkins Series in the Mathematical Sciences, The Johns Hopkins University Press*, (1992).

[11] W. Wang and K. Zhang. Colorings of hypergraphs. *Adv. Math. (China)*, **29(2)**, 115–136, (2000).

[12] D. B. West. Introduction to graph theory. *Prentice Hall*, 2nd ed., (2001).

November 21, 2004