THE TUKEY ORDER FOR GRAPHS

1. Introduction

Given a graph G, we let V_G stand for the vertex set of G. For graphs G without isolated vertices, the Tukey order can be characterized thus:

($*$) $G \leq H$ iff there exist functions $f : V_G \to V_H$ and $g : V_H \to V_G$ such that there is an edge from x to $g(y)$ every time there is an edge from $f(x)$ to y.

Any pair of functions (f, g) like this is called a *GT-connection* from G to H. If both $G \leq H$ and $H \leq G$, we write $G \approx H$. If $G \leq H$ but not $H \leq G$ we write $G < H$

The foregoing definition is worded so that it can be applied also to directed graphs in which each vertex has at least one edge going to it, and at least one edge emanating from it. [A more usual way of putting things is to say "... x is adjacent to $g(y)$ every time $f(x)$ is adjacent to y."] For more general directed graphs, including ordinary graphs with isolated vertices, the wording of the definition is different, and is usually given in terms of relations:

A binary relation R is a set (or also a proper class) of ordered pairs. The *domain of* R*, denoted* $dom\, R$, is the set $\{x : (\exists y)((x, y) \in R)\}$ and the *range of* R, denoted $ran\, R$, is $\{y : (\exists x)((x, y) \in R)\}$. The fact $(x, y) \in R$ is often written as xRy.

Let R and S be nonempty set relations. An ordered pair of functions (f, g) is called a *(generalized) Galois-Tukey connection* (abbreviated *GT-connection*) from R to S if the following holds:

(a) $f : dom\, R \longrightarrow dom\, S$

(b) $g : ran\, S \longrightarrow ranR$

(c) $(\forall x \in dom\, R)(\forall v \in ran\, S)(f(x), v) \in S$ implies $(x, g(v)) \in R$. We write $R \leq S$ if there is a GT-connection from R to S.

The canonical relation associated with a directed graph D has it that if x and y are in V_D then $(x, y) \in R(D)$ iff there is an edge from x to y. Ordinary graphs are thus associated with symmetrical relations.

This paper is concerned (except in this introduction and the final section) with ordinary graphs, and for these it is convenient in various statements to adhere to the modified definition ($*$) above. It is for graphs with isolated vertices that the two definitions give different results. In the Tukey order, isolated vertices are simply ignored in deciding whether $G \leq H$ since they are in neither $dom\, R(G)$ nor $ran\, R(G)$. In the modified order, any graph with an isolated vertex is $\geq$ every other graph. The directed graphs we consider in the last section will be ones for which the modified definition coincides with the Tukey definition.

2. Some basics

Even where graphs are concerned, the question of whether $G \le H$ promises in general to be NP-hard. For one thing, if $|G| = j$ and $|H| = k$ then there are $j^k \cdot k^j$ pairs of functions to investigate, while it only takes checking $(j^2 \cdot k^2)/4$ pairs of vertices to tell whether a given candidate is a GT-connection. For another, the existence of a GT-connection is often tied up with maximum and/or minimum sizes of certain kinds of subsets of graphs, as will be seen below, and finding these numbers is in some cases already known to be NP-hard. One of the few situations in which $G \le H$ is immediately obvious is the following:

Lemma 2.1. *Let G be a graph and let H be a subgraph of G such that $V_H = V_G$. Then $G \le H$.*

Indeed, the identity map on V_G works for both f and g in $(*)$. As will be seen later, this inequality need not be strict even if H is a proper subgraph of G with the same vertex set.

Definition 2.2. *A set A of vertices of a graph G is said to be* **bounded** *if there exists a vertex $z \in G$ such that there are edges from each $a \in A$ to z; otherwise A is* **unbounded**. *We say distinct vertices x, y of G are* **compatible** *if there exists $z \in G$ such that there are edges from both x and y to z. Otherwise x and y are* **incompatible**. *A subset A of V_G is* **pairwise [in]compatible** *if every pair of elements of A is [in]compatible. The* **incompatibility degree** *of G is the greatest size of a pairwise incompatible set of vertices.*

This use of the word "bounded" is taken from the theory of partially ordered sets, while our use of the word "compatible" is taken from set-theoretic forcing, where it is in turn inspired by logical properties of compatible elements in the forcing poset. The following definition is standard in graph theory.

Definition 2.3. *Let G be a graph without isolated vertices. A set C of vertices of G is* **open** *(or:* **total) dominating (in G)** *if for all vertices $x \in G$ there exists $c \in C$ such that there is an edge from x to c. The* **open** *(or:* **total) domination number** *of G is the least cardinality of an open dominating subset.*

For graphs without isolated vertices, it is easy to see that the incompatibility degree is never greater than the open domination number (which is undefined for graphs with one or more isolated vertices). The following is also easy to show:

Lemma 2.4. *The function f takes unbounded sets to unbounded sets and pairwise incompatible sets to pairwise incompatible sets; the function g takes open dominating subsets to open dominating subsets.*

We use the logicians' convention that $n = \{0, \dots n-1\}$ when n is a natural number. Let $(n, =)$ denote the graph with n vertices, each of which has a "loop on it" and no edges between different vertices. Clearly $(1, =) \le G$ for all graphs G. These graphs help analyze the GT-connections between graphs in general.

Theorem 2.5. *Let n be a positive integer and let G be a graph. Then $G \leq (n,=)$ iff the open domination number of G exists and is $\leq n$.*

Proof. ($\Longrightarrow$) Apply Lemma 2.4 to g.

($\Longleftarrow$) Let $\{x_0, \dots x_m\}$ be open dominating in G, $m < n$. For each vertex $x \in G$ let $f(x) = k$ iff there is an edge from x to x_k but not to x_j for $j < k$. Let $g(k) = x_k$ if $k < m$; otherwise any vertex of G will do for $g(k)$.

I suspect it is already known that finding the least cardinality of an open dominating set is NP-hard, and this would imply that just the problem of finding the least n such that $G \leq (n,=)$ is already NP-hard. Finding the greatest n such that $(n,=) \leq G$ also promises to be NP-hard, because of the next theorem.

Theorem 2.6. *Let n be a positive integer and let G be a graph. Then $(n,=) \leq G$ iff the incompatibility degree of G is at least n.*

Proof. ($\Longrightarrow$) Apply Lemma 2.4 to f.

($\Longleftarrow$) Let f take n one-to-one onto the members of an incompatible subset W of G. If $y \in G$ is a vertex then there is an edge from y to w for at most one $w \in W$. If there is such a w, let $w = f(k)$ and let $g(y) = k$; otherwise any member of n will do for $g(y)$.

Definition 2.7. *Let G be a graph without isolated vertices. The $\leq-$* **index** *of a graph G is the pair $\langle m, n\rangle$ where m is the incompatibility degree of G and n is its open domination number.*

Corollary 2.8. *If G has index $\langle m, n\rangle$ and H has index $\langle m', n'\rangle$ then:*
(1) If $n \leq m'$ then $G \leq H$.
(2) If $G \leq H$ then $m \leq m'$ and $n \leq n'$.

Corollary 2.9. *$G \approx (n,=)$ iff G has index $\langle n, n\rangle$.*

In contrast, if $1 \leq m < n$, there are infinitely many nonequivalent graphs with index $\langle m, n\rangle$, as we will see in later sections.

Despite the very different behavior (see 2.8) of graphs whose two indices are equal, there is a huge variety of graphs that are GT-equivalent to $(n,=)$ for any $n \geq 2$ because it is so easy to construct graphs with incompatible open dominating subsets. There are also many families of graphs in which the open domination number and incompatibility degree are the same, but are witnessed by different subsets.

Example 2.10. For each $n \geq 1$ let P_n be a simple path of length n. That is, P is a graph with vertices $v_0, v_1, \dots v_n$ and edges joining adjacent pairs v_i, v_{i+1}. Then P_4 has index $\langle 3, 3\rangle$; however, to get an open dominating family one must use the three middle vertices, while to get a pairwise incompatible family one must use two vertices on one end and one on the opposite end.

The indices for the P_n follow a regular pattern:
P_1, P_2 and P_3 all have index $\langle 2, 2\rangle$;
P_4 has index $\langle 3, 3\rangle$;
P_5, P_6 and P_7 all have index $\langle 4, 4\rangle$;
P_8 has index $\langle 5, 5\rangle$ .... In general,
if $k = 4n$, then the index of P_k is $\langle 2n+1, 2n+1\rangle$

if $k \in \{4n+1, 4n+2, 4n+3\}$ then the index of P_k is $\langle 2n+2, 2n+2\rangle$. In particular, each P_k is GT-equivalent to some $(n, =)$.

Problem 1. Is every tree GT-equivalent to $(n, =)$ for some n?

It is a challenge even for small graphs to find an efficient algorithm for determining whether two graphs are related by $<$ or $\approx$. For instance, it is relatively easy to show that a pentagon and a hexagon are incomparable in the order $\leq$, but it was tedious for me to show that a hexagon and a heptagon are also incomparable. On the other hand, it is an easy application of Theorems 2.8 and 2.9 that the hexagon (which has index $\langle 2, 4\rangle$) and the heptagon (which has index $\langle 3, 4\rangle$) are both strictly below the octagon (which has index $\langle 4, 4\rangle$). Another easy application is that the pentagon (which has index $\langle 2, 3\rangle$) is strictly below the heptagon.

Problem 2. Is the $4n + 2$-gon incomparable to both the $4n + 1$-gon and the $4n + 3$-gon for all positive integers n?

In other words, does the foregoing four-polygon pattern repeat endlessly? (It is easy to show that the $4n + 1$-gon has index $\langle 2n, 2n + 1\rangle$, the $4n + 2$-gon has index $\langle 2n, 2n+2\rangle$, the $4n+3$-gon has index $\langle 2n+1, 2n+2\rangle$, and the $4n+4$-gon has index $\langle 2n + 2, 2n + 2\rangle$. Then, arguing as for 5, 6, 7, and 8, we can confirm the remaining details of the pattern.)

3. Associated graphs and some of their applications

In this section we associate several different graphs with a given graph G. The following construction is especially useful if we define G by removing relatively few edges from K_n, the complete graph on n vertices, while keeping the vertices the same.

Construction 3.1. Let G be a graph. The *dual graph* $D(G)$ is defined as follows. $V_{D(G)=V_G}$, and vertices x, y of G (we allow $x = y$) are adjacent in $D(G)$ iff they are not adjacent in G.

Obviously, each graph is the dual of its dual: $D(D(G)) = G$. An easy proof by contrapositive shows:

Lemma 3.2. *$G \leq H$ iff $D(H) \leq D(G)$.*

In fact, (f, g) is a GT-connection from G to H iff (g, f) is a GT-connection from $D(H)$ to $D(G)$. In the language of category theory, we have a contravariant functor $\mathcal{D}$ from the category of graphs and GT-connections to itself, with $\mathcal{D}(G) = D(G)$ and $\mathcal{D}(f, g) = (g, f)$. Obviously, the composition of $\mathcal{D}$ with itself is the identity functor.

Example 3.3. Let K_n stand for the complete graph on the vertex set n. Then $D(K_n) = (n, =)$. If $m < n$ then it follows from Theorem 2.5 that $(m, =) < (n, =)$ and so Lemma 3.2 implies that $K_n < K_m$. Note also that K_1 (the trivial graph) is $\geq G$ for every graph G, that $K_2 \approx (2, =)$, and that for $n > 2$, the index of (K_n) is $\langle 1, 2\rangle$.

An interesting consequence of Lemma 3.2 and Theorems 2.5 and 2.6 is that the class of graphs that are not equivalent to either $(1, =)$ or K_1 has an infinite ascending cofinal sequence —the graphs $(n, =)$—and an infinite descending coinitial sequence, the graphs K_n. Putting these two sequences together, we have:

$(n > 3)$

$$(1,=) < \cdots < K_n < \cdots < K_3 < K_2 \approx (2,=) < (3,=) < \cdots < (n,=) < \cdots < K_1$$

The following corollary is an easy consequence of Lemma 3.2 and Theorems 2.5 and 2.6.

Corollary 3.4. *Let G be a graph. Then $G \leq K_n$ iff $D(G)$ has a pairwise incompatible subset of cardinality $\geq n$, and $K_n \leq G$ iff $D(G)$ has an open dominating subset of cardinality $\leq n$.*

When G is a simple graph, every vertex of $D(G)$ has a loop on it. To better analyze such graphs, we introduce two more constructions.

Definition 3.5. *Let G be a graph.* The augmented graph of G, *denoted $A(G)$, is formed by adding a loop to every vertex in G that doesn't already have one. The* simple graph associated with G, *denoted $S(G)$, is G with all loops omitted.*

The index of a graph of the form $A(G)$ can be characterized in familiar terms, using the following well-known concepts of graph theory:

Definition 3.6. *A set S of vertices of a graph G is said to be* **dominating** *if every vertex of G is either in S or adjacent to some vertex in S, and to be* **a packing** *if all its members are at a distance of ≥ 3 from each other.*

Lemma 3.7. *A set S of vertices of the graph G is pairwise incompatible in $A(G)$ iff it is a packing in $A(G)$ iff it is a packing in G iff it is a packing in $S(G)$.*

Proof. The first "iff" is true because in $A(G)$, vertices at a distance ≤ 2 are compatible, and the others are true because packings are unaffected by the presence or absence of loops.

Lemma 3.8. *A set S of vertices of the graph G is open dominating in $A(G)$ iff it is dominating in $A(G)$ iff it is dominating in G iff it is dominating in $S(G)$.*

The first "iff" is true because each vertex is adjacent to itself in $A(G)$ while the second is true because domination is unaffected by loops.

For simple graphs, $A(D(G)) = D(G)$, and thus the indices of the dual graphs can be found by finding their packing number and their dominating number. Then (see Corollary 2.8 and Lemma 3.2) if the index of $D(G)$ is not the index of $D(H)$, G is not equivalent to H. Also, if $D(G)$ and $D(H)$ have dominating number coinciding with packing number, and these numbers are the same for $D(G)$ and $D(H)$, then we know that G is GT-equivalent to H by Corollary 4.5.

Example 3.9. Every star has a dominating packing of size 1, consisting of the one vertex which is adjacent to all the others. If we remove the edges of a star X with k edges from K_n to get G, then $D(G)$ is the direct sum of $n-k-1$ copies of $(1,=)$ and one copy of $A(X)$. Hence $D(G)$ is equivalent to $(n-k,=)$ and so is $D(K_{n-k})$, so G is equivalent to K_{n-k}.

If we also remove from K_n any number of edges connecting the vertices of X, and let H be the resulting graph, then $D(H)$ is still equivalent to $(n-k,=)$ because adding edges to X does not affect its dominant packing. So H is also equivalent to K_{n-k}.

Example 3.10. As is well known, every path has a packing which is also dominating. Hence if G is obtained from K_n by removing the edges of a path, then G is GT-equivalent to a complete graph. To determine which one it is, say the path is $P_k (0 < k \leq n)$. Let $k = 3j + i$, where $i \in \{0, 1, 2\}$; then $D(G)$ is the direct sum of $n - k - 1$ copies of $(1, =)$ and one copy of $A(P_k)$, which has a dominating packing of size $j + 1$. Hence $G \approx K_{n-k+j}$.

Example 3.11. Let G, H and K be hexagons with vertices numbered cyclically in order 1, 2, 3, 4, 5, 6. Besides the usual edges, let G include the edges of the internal triangle with vertices 1, 3, and 5; let H have edges connecting 1 with every other vertex; and let K have edges connecting 1 with 3 and 5, and 4 with 2 and 6. It is easy to see that all three graphs are of index $\langle 1, 2\rangle$. Their duals are simpler than the graphs themselves, and (see below) $D(H)$ and K have indices $\langle 2, 3\rangle$, while $D(G)$ has index $\langle 3, 3\rangle$ (see below), so $G < H$ and $G < K$.

Also, $H \approx K$ as the following analysis of the dual graphs shows. Besides the ubiquitous loops, $D(H)$ has a pentagon with one diagonal added, plus one isolated point; while $S(D(K))$ is the direct sum of a graph with vertices $\{1, 4\}$ connected by an edge and a rectangle with vertices $\{2, 5, 3, 6\}$ arranged cyclically in that order. The pentagon of $D(G)$ has vertices $\{2, 5, 3, 6, 4\}$ arranged cyclically in that order, and a diagonal connecting 2 and 6. Now let $f : V_H \to V_K$ satisfy $f(4) = 6$ and be the identity on the other vertices, while $g : V_K \to V_H$ satisfies $g(4) = 1$ and is otherwise the identity. This shows $D(H) \leq D(K)$ and so $K \leq H$. Now (g, f) is a GT-connection from $D(K)$ to $D(H)$ and so $H \leq K$ also.

As for G, its dual has a triangle with edges attached to each corner. This has a dominant packing consisting of the far ends of the edges. So $D(G)$ has index $\langle 3, 3, \rangle$, making it equivalent to K_3. If we add diagonals to G, the effect on $D(G)$ is to erase the edges that come out of the vertices of the triangle. Erasing one edge leaves a connected graph in which there is a packing of size 2 and a dominating set of size 2; they cannot coincide, but the resulting dual graph still has index $\langle 3, 3, \rangle$. For similar reasons, we can also add two or all three diagonals to G and still have the resulting graph equivalent to K_3.

4. More associated graphs and a dual index

The associated graphs in the preceding section shared the same vertex set as the original graph. The ones in this section may have a different vertex set.

Construction 4.1. *Let G be a graph of incompatiblity degree ≥ 2. The* **incompatibility graph of G,** *denoted $I(G)$, is the graph whose edges are incompatible pairs $\{x, y\}$ of vertices of G and whose vertices are the union of the edges.*

The following is immediate from Lemma 2.4.

Lemma 4.2. *If (f, g) is a GT-connection from G to H then $f \restriction V_{I(G)}$ is an edge-preserving map from $I(G)$ to $I(H)$.*

Example 4.3. Let K be a pentagon and let H and G be hexagons with vertices numbered cyclically in order 1, 2, 3, 4, 5, 6, and with one diagonal [*resp.* two diagonals] connecting vertices 6 and 4 [*resp.* and 5 and 3]. All three graphs are of index $\langle 2, 3\rangle$ and their duals all are of index $\langle, 1, 2\rangle$ so the dual graphs are not immediately helpful. Clearly (see Lemma 2.1) $G \leq H$, and $G \leq K$ because G is formed by replacing one vertex of K by a copy of K_2 whose vertices are adjacent to exactly the same vertices as is the one vertex they replace.

These inequalities are strict, and the incompatibility graphs give an easy avenue to proving$G < K$ and $K \not\leq H$. The incompatibility graph of the pentagon K is simply K again; $I(G)$ is a path of length 3 connecting the vertices 6, 1, 2, and 3 in that order; and $I(H)$ adds 4 and 5 to its vertex set, with 4 at the end of the path and an edge joining 2 to 5. A simple parity argument shows that a polygon with an odd number of sides does not admit of an edge-preserving map to a path, hence not to a tree like $I(H)$ either. So $G < K$ and $K \not\leq H$. We will show $H \not\leq K$, thereby also establishing $G < H$.

All three graphs share the following property: any two compatible vertices have *only one* adjacent vertex in common. Now for any GT-connection (f, g) from H to K, there must be a pair of vertices $\{v_1, v_2\}$ of H sent to the same vertex of K by f. In K, there are two vertices, w_1, w_2 adjacent to $f(v_1)$, and $\{f(v_1), w_1, w_2\}$ is open dominating in K, but g would have to take w_1 and w_2 to the (at most!) one vertex that is adjacent to both v_1 and v_2, but this contradicts Lemma 2.4 and the fact that the index of H is $\langle 2, 3\rangle$.

Corollary 3.4 and examples like 3.11 make it natural to define a dual index for graphs. To help avoid confusion, we make the numbers negative; this is partly motivated by the fact that the sequence displayed in the preceding section is of order type $\mathbb{Z}^+ \cup \{-\infty, +\infty\}$ with the complete graphs of cardinality greater than 2 playing the role of the negative numbers.

Definition 4.4. *The* **dual index** *of a graph G is $\langle -m, -n\rangle$ where $\langle m, n\rangle$ is the index of $D(G)$.*

The following is almost immediate from Corollary 2.8.

Corollary 4.5. *If G has dual index $\langle -m, -n\rangle$ and H has dual index $\langle -m', -n'\rangle$ then:*
(1) If $n \leq m'$ then $H \leq G$.
(2) If $H \leq G$ then $m \leq m'$ and $n \leq n'$.

The second coordinate of the dual index has a direct characterization in terms of G: its absolute value is the least cardinality of an unbounded set of vertices of G. This is immediate from the following easy lemma:

Lemma 4.6. *Let G be a graph and let $B \subset V_G$. Then B is unbounded in G iff it is dominating in $D(G)$.*

Theorem 4.7. *Let G be a graph with index $\langle m, n\rangle$ and dual index $\langle -j, -k\rangle$. Exactly one of the following is true:*
(1) $m = 1$ *and* $k > 2$
(2a) $m = n = j = k = 2$
(2b) $m = 2, n > 2, j = 1,$ *and* $k = 2$
(3) $m > 2, j = 1$ *and* $k = 2$

Proof. If $m = 1$, then the smallest unbounded subset of G has more than two elements, so $k > 2$. If $m \geq 2$ then $k \leq 2$ since every incompatible pair is unbounded; but then $k = 2$, because otherwise $D(G) \approx (1, =)$ and G has an isolated vertex and its index is undefined. Similarly if $j = 2$ then $n = 2$. This and the inequalities

$j \leq k$ and $m \leq n$ account for both $(2a)$ and $(2b)$. If $m > 2$ then $G \geq (3, =)$ so $D(G) \leq K_3$, and K_3 has index $\langle 1, 2\rangle$. Since G has an index, the index of $D(G)$ cannot be $\langle 1, 1\rangle$ so it must be $\langle 1, 2\rangle$.

Construction 4.8. *Let G be a graph with index $\langle 1, 2\rangle$ or $\langle 2, 2\rangle$. The* open domination graph *of G, denoted $O(G)$, is the graph whose edges are open dominating subsets of G and whose vertices are the union of the edges of G.*

Example 4.9. Let $(n, \square)$ be the n-gon with vertex set n arranged cyclically. Let G_n be a graph obtained by removing the edges of an n-gon from K_m, $m \geq n$. Then $D(G)$ is naturally isomorphic to $A(n, \square)$, whose index can be characterized as follows: $A(3n, \square)$ has index $\langle n, n\rangle$; and $A(3n+1, \square)$ and $A(3n+2, \square)$ both have index $\langle n, n+1\rangle$. It is easy to see that $A(n, \square) \leq A(m, \square)$ whenever $n \leq m$, and that the inequality is strict if m or n is $3k$ for some integer k: the latter fact follows from the former and Corollary 2.9, while the former can be established by having f be the inclusion map while g takes $\{n-1, \ldots m\}$ to $n-1$ and is otherwise the identity.

We also have that $A(4, \square) < A(5, \square)$ as a glance at the open domination graphs shows. Any two distinct vertices of $A(4, \square)$ give an edge of $O(A(4, \square))$, which is thus a copy of K_4 while $O(A(5, \square))$ is a pentagon. There can be no edge-preserving map from K_4 (or indeed from any graph containing a triangle) to a pentagon, so the demands placed on g by Lemma 2.4 on any GT-connection from $A(5, \square)$ to $A(4, \square)$ cannot be met.

5. Simple graphs of index $\langle 1, 2\rangle$