Strongly connectable digraphs and non-transitive dice

Simon Joyce[*], Alex Schaefer[†], Douglas B. West[‡§], and Thomas Zaslavsky[¶]

July 31, 2015

Abstract

We show that a directed graph extends to a strongly connected digraph on the same vertex set if and only if it has no complete directed cut. We bound the number of edges needed for such an extension, and we apply the characterization to a problem on non-transitive dice.

1 Introduction

Characterization theorems in graph theory are important because they often guarantee short certificates for both a "yes" and "no" answer to the question of whether a graph satisfies a particular property. Many such theorems state that obvious necessary conditions are also sufficient. In this note we prove such a result about directed graphs and apply it to a problem on non-transitive dice from [12].

A *strict digraph* is a directed graph in which each unordered pair of vertices is the set of endpoints of at most one edge; that is, a strict digraph is an orientation of a simple graph. A digraph is *strongly connected* if it contains a (directed) path from x to y for every ordered pair of vertices. A digraph G *extends* to a digraph G' if $V(G) = V(G')$ and $E(G) \subseteq E(G')$. We ask when a (strict) digraph extends to a strongly connected strict digraph. Note that it does so if and only if it extends to a strongly connected tournament, where a *tournament* is an orientation of a complete graph.

[*]Binghamton University (SUNY), joyce@math.binghamton.edu

[†]Binghamton University (SUNY), schaefer@math.binghamton.edu

[‡]Departments of Mathematics, Zhejiang Normal University & University of Illinois, dwest@math.uiuc.edu.

[§]Research supported by Recruitment Program of Foreign Experts, 1000 Talent Plan, State Administration of Foreign Experts Affairs, China.

[¶]Binghamton University (SUNY), zaslav@math.binghamton.edu

For a set $X \subseteq V(G)$, let $\overline{X} = V(G) - X$. When $\varnothing \neq X \subset V(G)$, the *cut* $[X, \overline{X}]$ is the set $\{xy \in E(G)\colon\ x \in X, y \in \overline{X}\}$, where we write xy for an edge with tail x and head y. A cut $[X, \overline{X}]$ is a *dicut* if there is no "back edge" of the form yx with $y \in \overline{X}$ and $x \in X$. It is a *complete dicut* if it contains all $|X| \cdot |\overline{X}|$ edges of the form xy with $x \in X$ and $y \in \overline{X}$.

Obviously, a digraph that extends to a strongly connected strict digraph contains no complete dicut. This obvious necessary condition is also sufficient.

Theorem 1 (Strong connectability). *A strict digraph extends to a strong strict digraph if and only if it contains no complete dicut.*

We call such a digraph *strongly connectable.* We prove sufficiency in Section 2 from a stronger result, Theorem 2, giving an upper bound on how many edges need to be added.

The problem of deciding strong connectability belongs to the class NP $\cap$ co-NP. A certificate for strong connectability is a strongly connected extension, and a certificate for not being strongly connectable is a complete dicut. Both are easy to confirm. The latter is obvious. For the former, there exists an extension to a strong tournament, which contains a spanning cycle by [8, Section 4].

Section 3 applies Theorem 1 to a problem on "non-transitive dice" discussed in Schaefer [12]. A set of dice is *non-transitive* if there is a cycle such that each die beats the next in cyclic order, and it is *balanced* if there exists a value $p > 1/2$ such that for any two dice, one beats the other with probability exactly p. Schaefer asked which digraphs are realizable by balanced non-transitive dice; he showed that those digraphs are precisely the ones that are strongly connectable (a result that led to our investigation of strong connectability). Thus our theorem provides a criterion for a digraph to be realizable by balanced non-transitive dice, where a set of dice *realizes* a digraph if the vertices can be assigned to dice in the set so that when uv is an edge, the die representing u beats the die representing v.

An analogue of our problem was studied earlier. Eswaran and Tarjan [1] studied strong connectability for general digraphs, which allow antiparallel pairs of edges. This makes a huge difference. Without strictness, every digraph extends to a strong digraph, simply by introducing every ordered pair of distinct vertices as an edge. Thus, their problem becomes that of finding the minimum number of edges to add and an algorithm to produce a smallest strong extension. Frank [2] and Frank and Jordán [3, 4] generalized the question to extensions that have connectivity or edge-connectivity at least k, again allowing antiparallel edge pairs and again seeking the smallest such extension and an algorithm.

The problem for strict digraphs is very different, beginning with the fact that not every strict digraph is strongly connectable. We also do not give an exact minimum number or set of edges. The complexity of determining the minimum size of an extension set remains open, and it seems hard to generalize Theorem 1 to characterize digraphs extendable to a strict k-connected digraph.

2 Strongly connectable digraphs

A *strong component* of a digraph G is a maximal strongly connected subgraph. The strong components of G form an acyclic digraph called the *condensation* G^*, obtained by contracting each strong component to a vertex and eliminating duplicate edges. A strong component is a *source* or *sink component* of G according as it is a source or sink vertex in G^* (it may be both). A *successor* of a strong component C is another strong component C', such that there is a path from C to C' in G'; when C' is a successor of C, also C is a *predecessor* of C'. For $v \in V(G)$, let $C(v)$ denote the strong component of G containing v. The *underlying graph* of a digraph G is the graph $\hat{G}$ obtained by treating its edges as unordered pairs. We say that G is *weakly connected* when $\hat{G}$ is connected. The *weak components* of a digraph G are the subdigraphs induced by the vertex sets of components of $\hat{G}$.

We prove Theorem 1 by obtaining an upper bound on the number of edges needed. We will strengthen the upper bound for disconnected graphs in Lemma 4 and Proposition 5.

Theorem 2 (Upper bound)**.** *Let G be a strict digraph having at least three vertices and r strong components. If G has no complete dicut, then G extends to a strongly connected strict digraph by adding at most r edges, with equality if and only if $\hat{G}$ is disconnected and each weak component of G is strong.*

Proof. Suppose that $\hat{G}$ is disconnected and every weak component of G is strong. If $r > 2$, then choose one vertex from each component and add a cycle of r edges through them to obtain a strongly connected extension. If $r = 2$, then since G has at least three vertices, one weak component has at least two vertices, and we can add edges to and from distinct vertices in that component to obtain a strong extension. Since every weak component must receive an added entering edge, equality holds in this case.

Henceforth we may assume that not every weak component is strong, so G has a strong component that is a sink component but not a source component.

Case 1: *G is weakly connected.* To prove the upper bound $r - 1$, we use induction on r. For $r = 1$ there is nothing to prove (no edge need be added). Consider $r > 1$.

Let S be the set of all vertices in source components of G. Since $[S, \overline{S}]$ is a dicut, there is a pair (y, x) with $y \in S$ and $x \in \overline{S}$ such that $yx \notin E(G)$. Add edge xy to G, forming a new strict digraph G'. If $C(x)$ is a successor to $C(y)$ in G, then G' has a strong component containing $C(x)$ and $C(y)$, so G' has at most $r - 1$ strong components. Otherwise, $x \notin S$ implies that $C(x)$ has some source component as a predecessor; let z be a vertex in that component, and let $G'' = G' + yz$. Since no edge connects two source components, G'' is a strict digraph. It has at most $r - 2$ strong components, since $C(x)$, $C(y)$, and $C(z)$ all lie in a single strong component of G''. It now suffices by the induction hypothesis to show that G' and in the second case also G'' has no complete dicut.

Let $[X, Y]$ be a complete dicut in G'. As G has no complete dicut, the added edge xy must satisfy $x \in X$ and $y \in Y$. Thus $C(x) \subseteq X$ and $C(y) \subseteq Y$. In G no edges enter the source component $C(y)$, so in G' only one edge enters $C(y)$. Since $C(y) \subseteq Y$, this implies $|V(C(y))| = 1$. That makes y a source vertex in G, so only the edge xy enters it in G'. Thus $|X| = 1$. This implies that x is a source vertex in G' and therefore in G, which contradicts $x \notin S$. We conclude that G' has no complete dicut.

In the case where $C(x)$ is not a successor to $C(y)$, the source component $C(z)$ in G remains a source component in G'. Also, the strong component of G' that contains both $C(x)$ and $C(y)$ from G is a successor of $C(z)$ in G'. Thus G'' is formed from G' by adding yz in the way that G' was formed from G by adding xy. Since G' satisfies the same hypotheses required of G, the same argument now implies that G'' also has no complete dicut.

Case 2: *G is not weakly connected.* Let $G_1, \ldots, G_k$ with $k > 1$ be the weak components of G. In each G_i choose a source component S_i and a sink component T_i such that $T_i = S_i$ if G_i is strong and otherwise T_i is a successor of S_i. Treating subscripts modulo k, for $1 \leq i \leq k$ add an edge $t_i s_{i+1}$ such that $t_i \in V(T_i)$ and $s_{i+1} \in V(S_{i+1})$.

The resulting digraph G' is weakly connected. It is obviously strict when $k > 2$, and when $k = 2$ the edges $t_2 s_1$ and $t_1 s_2$ do not have the same pair of endpoints because at least one weak component is not strong. For the same reason, the k added edges are too few to complete a complete dicut unless $k = 2$ and G has exactly three vertices, but then G' is a (directed) cycle. Thus, the connected case applies to G'.

Let r' be the number of strong components of G'. Since all S_i and all T_i lie in one strong component in G', and there are at least $k + 1$ such sets since $S_i = T_i$ only when G_i is strong, we have $r' \leq r - k$. By Case 1, G' can be made strongly connected by adding at most $r' - 1$ edges, so the number of edges needed to make G strongly connected is at most $r - 1$. □

Remark 3. Since every source component needs an entering edge and every sink component needs an exiting edge, at least $\max(s, t)$ added edges are needed for a strongly connected extension, where s and t are the numbers of source and sink components.

Example 1. The upper bound $r - 1$ is sharp for weakly connected strict digraphs. Consider the digraph obtained from a transitive tournament with r vertices by deleting the unique spanning path. There are r strong components and no complete dicut, and extending to a strong strict digraph requires adding the $r - 1$ missing edges. When $r \geq 4$, this example has two source components and two sink components, so the lower bound in Remark 3 can be arbitrarily bad when G is weakly connected.

Example 2. The upper bound $r - 1$ is also sharp for digraphs with more than one weak component when the components are not all strong. Let G be the digraph formed from the complete bipartite graph $K_{p,q}$ with bipartition (X, Y), where $|X| = p$ and $|Y| = q$, by directing each edge from X to Y and adding an isolated vertex. A strong extension must add edges entering each vertex in X and edges leaving each vertex in Y, but no edge can do both. Thus, the needed number of added edges is at least $p + q$, which equals $r - 1$. Here again the bound of Remark 3 is weak.

In spite of Example 2, the upper bound can be reduced in the disconnected case by using additional information, such as the number of weak components that are not strong, the number of strong components that are neither sources nor sinks, or the number of source and sink components in each weak component. We omit the details of the first two; the next result concerns the last.

Lemma 4 (Disconnected upper bound)**.** *Let G be a strict digraph that is not weakly connected. Let $G_1, \ldots, G_k$ be its weak components, and let s_i and t_i be the number of source and sink components, respectively, in G_i. The minimum number of edges that must be added to make G strongly connected is at most* $\max(t_1, s_2) + \cdots + \max(t_{k-1}, s_k) + \max(t_k, s_1)$.

Proof. We add edges from sink components of G_{i-1} to source components of G_i, viewing subscripts modulo k. We add at least one edge leaving each sink component and one edge entering each source component. This is easy to do using $\max(t_{i-1}, s_i)$ edges. The resulting digraph is obviously strongly connected and, when $k \geq 3$, strict. When $k = 2$, the same observation as in Case 2 of Theorem 2 allows it to be strict. □

The exact value of this upper bound depends on the cyclic order chosen for the components of G. We do not know a procedure to minimize the sum. The lower bound from Remark 3 is $\max(\sum t_i, \sum s_i)$; it equals the upper bound from Lemma 4 when the comparison of t_{i-1} and s_i goes the same way for each i. Hence both bounds are sharp.

Recall that the weak components of G correspond to the components of $\hat{G}$.

Proposition 5 (Disconnected upper bound)**.** *A strict digraph G that is not weakly connected can be strongly connected by adding at most $s+t-c$ edges, where G has s source components, t sink components, and c weak components. The bound equals $u-c'$, where u strong components are source or sink components and c' weak components are not strong components.*

Proof. This follows from Lemma 4, since $\max(t_{i-1}, s_i) \leq t_{i-1}+s_i-1$ (again taking subscripts modulo k), giving an upper bound of $\sum t_{i-1} + \sum s_i - c$. The sum is $s + t - c$. The sum $s + t$ exceeds u by the number $c - c'$ of weak components that are strongly connected; hence $s + t - c = u - c'$. □

Proposition 5 strengthens the upper bound in Theorem 2 when G is not weakly connected. By definition, always $u \leq r$, so the upper bound of r is reduced by at least 1 for each weak component that is not strong.

Example 3. Toward understanding the problem of obtaining strong extensions by adding the fewest edges, it is interesting to consider digraphs obtained from bipartite graphs. Let G be a strict digraph obtained from a bipartite graph with bipartition (X, Y) by orienting all edges from X to Y. We forbid the underlying graph to be complete bipartite, because the orientation would yield a complete dicut. Let $s = |X|$ and $t = |Y|$.

As in Remark 3, a strongly connected extension of G must always add at least $\max(s, t)$ edges; this is the trivial lower bound. By symmetry, suppose $s \geq t$. We claim that achieving this lower bound requires adding edges not in $\hat{G}$ that match Y into X. To see this, note that s added edges must enter X and t added edges must exit Y. To do this using only s edges, each of the t added edges leaving Y must be one of the s edges entering X. These t edges form a matching of Y into X.

If the digraph on $2t$ vertices induced by the vertices covered by this matching is strongly connected, then adding an edge from the matched vertices of X to each of its $s-t$ remaining vertices completes the desired strongly connected extension.

At the other extreme, when G has $st - 1$ edges, only the one edge missing from $K_{s,t}$ can connect a sink to a source, and the upper bound $s + t - 1$ cannot be improved.

The common generalization of the two extremes improves the lower bound to $s + t - m$, where m is the maximum size of a matching from Y into X using edges not in the original bipartite graph.

A digraph G is *k-connected* if it has more than k vertices and any deletion of fewer than k vertices from G leaves a strong digraph. For a generalization analogous to those studied for non-strict digraphs, we say that a strict digraph G is *k-connectable* if it extends to a k-connected strict digraph. An obvious necessary condition for k-connectability is that every dicut $[X, \overline{X}]$ lacks at least k edges. It is not clear whether this condition is sufficient; our proofs for $k = 1$ do not extend, because when $k \geq 2$ the maximal k-connected subgraphs of a graph need not be pairwise disjoint.

3 Non-transitive dice

In a set of ordinary dice, all dice are the same and the probability that one die rolls a higher number than another is exactly 1/2. Martin Gardner publicized the idea, due to Bradley Efron, of dice where not only is the probability other than 1/2, but there can be three dice such that each beats one of the others with probability greater than 1/2 (see [5, 6, 7]). Such dice are *non-transitive*: generalizing to n dice, there exist l that can be arranged cyclically so that each has probability greater than 1/2 of rolling higher than its successor.

Quimby found a set of four non-transitive 6-sided dice whose 24 sides are the distinct numbers from 1 to 24 [9] (see [10] for other sets of non-transitive dice). Schaefer and Schweig [11] carried the idea further; they studied k-sided generalized dice in which each die has k different numbers on it and the numbers on all dice are distinct. We may regard each die as a set of k distinct integers and define a *set of dice* D as a set of pairwise disjoint such sets. As they observed, one can always choose the dice to partition the set $\{1, \ldots, kn\}$.)

We say that die D_1 *beats* D_2 and write $D_1 \succ D_2$ if, among all pairs of numbers in $D_1 \times D_2$, the first number is larger than the second more than half the time. A set of dice is *transitive* if the relation $\succ$ is transitive. Although it may be contrary to intuition, a randomly chosen set of (more than two) dice need not be transitive. Schaefer and Schweig [11] showed that it is easy to make an intransitive set of three or four k-sided dice when $k \geq 3$.

The relation $\succ$ can be represented by a digraph $G(D)$ with one vertex for each die and an edge from i to j if D_j beats D_i. The relation $\succ$ is antisymmetric, meaning that $D \succ D'$ and $D' \succ D$ imply $D = D'$, so the digraph $G(D)$ is strict. If H is any subgraph of $G(D)$, we say that D *realizes* H; that means every edge of H corresponds to a pair of dice in D in which one beats the other as indicated by the direction of the edge. (H need not be an induced subgraph.)

Let $p_{i,j}$ be the probability that D_i beats D_j; that is, $p_{i,j}$ is the proportion of pairs in $D_i \times D_j$ in which the first number is the larger. (Because no two numbers on dice are equal, $p_{i,j} + p_{j,i} = 1$.) Schaefer and Schweig [11] call a set of dice *balanced* if all unordered pairs $\{p_{i,j}, p_{j,i}\}$ are the same. They found that for $n = 3 \leq k$ it is possible to form non-transitive dice that are balanced. Schaefer [12] then showed that for a tournament T of order $n \geq 3$, there is a balanced set of n non-transitive dice that realizes T if and only if T is strongly connected. He observed the corollary that non-transitive dice realizing a strict digraph G can be chosen balanced if and only if G is strongly connectable. Thus our main theorem gives a criterion for the existence of balanced dice realizing a given relation (transitive or not).

Corollary 6. *An antisymmetric relation is realizable by a set of balanced dice if and only if its digraph has no complete dicut.*

References

[1] Kapali P. Eswaran and Robert E. Tarjan, Augmentation problems. SIAM J. Comput. 5(4) (1976), 653–665.

[2] András Frank, Augmenting graphs to meet edge-connectivity requirements. SIAM J. Discrete Math. 5(1) (1992), 25–53.

[3] András Frank and Tibor Jordán, Minimal edge-coverings of pairs of sets. J. Combin. Theory Ser. B 65(1) (1995), 73–110.

[4] András Frank and Tibor Jordán, Directed vertex-connectivity augmentation. In *Connectivity augmentation of networks: structures and algorithms (Budapest, 1994)*. Math. Program. 84B(3) (1999), 537–553.

[5] Martin Gardner, The paradox of the nontransitive dice and the elusive principle of indifference. Scientific American 223(12) (Dec., 1970), 110–114.

[6] Martin Gardner, On the paradoxical situations that arise from nontransitive relations. Scientific American 231(10) (Oct., 1974), 120–125.

[7] Martin Gardner, Nontransitive dice and other paradoxes. Ch. 22 in Martin Gardner, The Colossal Book of Mathematics: Classic Puzzles, Paradoxes, and Problems, pp. 286–296. Norton, New York, 2001.

[8] John W. Moon, Topics on Tournaments. Holt, Rinehart and Winston, New York, 1968.

[9] Shirley Quimby, in The Pallbearers Review (Jan., 1971).

[10] Richard P. Savage Jr., The paradox of nontransitive dice. Amer. Math. Monthly 101(5) (May, 1994), 429–436.

[11] Alex Schaefer and Jay Schweig, Balanced non-transitive dice. Submitted.

[12] Alex Schaefer, Balanced non-transitive dice, II. In preparation.