

Reverse mathematics of some topics from algorithmic graph theory

by

Peter G. Clote (Chestnut Hill, Mass.) and
Jeffrey L. Hirst (Boone, N.C.)

Abstract. This paper analyzes the proof-theoretic strength of an infinite version of several theorems from algorithmic graph theory. In particular, theorems on reachability matrices, shortest path matrices, topological sorting, and minimal spanning trees are considered.

For finite graphs, polynomial time algorithms can be used to find reachability matrices, shortest path matrices, topological sortings and minimal spanning trees. We will analyze the proof-theoretic and recursion-theoretic strength of infinite statements related to these topics using the techniques of reverse mathematics.

The subsystems of reverse mathematics which are used below include \mathbf{RCA}_0 , \mathbf{ACA}_0 , and $\mathbf{\Pi}_1^1\text{-CA}_0$. The system \mathbf{RCA}_0 consists of basic arithmetic axioms, a restricted induction scheme, and a comprehension axiom asserting the existence of Δ_1^0 definable sets. \mathbf{ACA}_0 adds to \mathbf{RCA}_0 a comprehension axiom for arithmetically definable sets. Similarly, $\mathbf{\Pi}_1^1\text{-CA}_0$ includes a comprehension axiom for Π_1^1 definable sets. A nice introduction to reverse mathematics can be found in [4]. Details on the encoding of graph-theoretic statements in subsystems of second order arithmetic can be found in [1].

Reachability and shortest path matrices. Given a graph $G = (V, E)$ with vertex set V and edge set E , the *reachability matrix* for G is a function $M : V \times V \rightarrow \{0, 1\}$ such that $M(v, w) = 1$ if and only if there is a path in G from v to w . The next three results explore the proof-theoretic and

1991 *Mathematics Subject Classification*: 03F35, 03D45.

Key words and phrases: reverse mathematics, proof theory, recursion theory, graph theory.

Clote's research was partially supported by NSF grant DCR8606165.

[1]

recursion-theoretic strength of statements concerning reachability matrices. The following theorem is closely related to Theorem 2.5 of [2]. Clauses in parentheses may be included or deleted without affecting the validity of the theorem. A graph is *2-regular* if every vertex is contained in exactly two edges.

THEOREM 1. (\mathbf{RCA}_0) *The following are equivalent:*

(1) \mathbf{ACA}_0 .

(2) *If G is a (2-regular) (directed) graph, then the reachability matrix of G exists.*

Proof. To prove that (1) implies (2), assume \mathbf{ACA}_0 and let G be a graph. The reachability matrix of G can be defined by setting $M(v, w) = 1$ if and only if some finite sequence of edges forms a (directed) path from v to w . Since the non-zero entries of M are a Σ_1^0 definable set, by \mathbf{ACA}_0 the reachability matrix exists.

To prove that (2) implies (1), assume \mathbf{RCA}_0 and the version of (2) which includes all the parenthetical hypotheses. By Lemma 2.3 of [3] it suffices to prove the existence of the range of an arbitrary injection, denoted by $f : \mathbb{N} \rightarrow \mathbb{N}$. Define the graph G as follows. The vertices of G will be of the form l_i^j and r_i^j for all $i, j \in \mathbb{N}$. Include all edges of the form (r_i^0, l_i^0) in G . For each i and j , if $f(j) \neq i$, then include the directed edges (l_i^j, l_i^{j+1}) and (r_i^{j+1}, r_i^j) in G . On the other hand, whenever $f(j) = i$, include the edges (l_i^j, r_i^j) and (l_i^{j+1}, r_i^{j+1}) in G . Since G is Δ_1^0 definable, \mathbf{RCA}_0 proves its existence. G is a 2-regular directed graph, so (2) implies that the reachability matrix of G exists. Since for each $i \in \mathbb{N}$, i is in the range of f if and only if there is a directed path from l_i^0 to r_i^0 , the range of f is Δ_1^0 definable in the reachability matrix of G . Thus, by \mathbf{RCA}_0 , the range of f exists, completing the proof. ■

COROLLARY 2. (\mathbf{RCA}_0) *The following are equivalent:*

(1) \mathbf{ACA}_0 .

(2) *For any sequence $\langle G_i \mid i \in \mathbb{N} \rangle$ of (2-regular) (directed) graphs, there is a sequence $\langle M_i \mid i \in \mathbb{N} \rangle$ such that for each i , M_i is the reachability matrix for G_i .*

Proof. To prove that (1) implies (2), it suffices to note that the sequence of non-zero entries of the reachability matrices is Σ_1^0 definable in the sequence of graphs. The proof that (2) implies (1) follows immediately from Theorem 1. ■

COROLLARY 3. *There is a recursive (2-regular) (directed) graph G such that $0'$ is recursive in the reachability matrix of G . If $\langle G_i \mid i \in \mathbb{N} \rangle$ is a recur-*

sive sequence of (2-regular) (directed) graphs, then the sequence of reachability matrices for these graphs is recursive in $0'$.

Proof. To prove the first statement, imitate the proof of Theorem 1, letting f be a recursive function such that $0'$ is recursive in the range of f . The second statement follows immediately from the Σ_1^0 definability argument in the proof of Corollary 2. ■

Additional information can be encoded in the reachability matrix without changing the proof-theoretic strength of the statement. In the following, we refer to acyclic paths as *arcs*. Also, we say that two arcs $v, a_0, a_1, \dots, a_m, w$ and $v, b_0, b_1, \dots, b_n, w$ are *distinct* if for some j , $a_j \neq b_j$.

THEOREM 4. (\mathbf{RCA}_0) *The following are equivalent:*

(1) \mathbf{ACA}_0 .

(2) *If G is a (2-regular) (directed) graph, then there is a function $M : V \times V \rightarrow \mathbb{N} \cup \{\infty\}$ such that for all pairs of vertices v and w of G , $M(v, w)$ is the number of distinct arcs from v to w .*

Proof. To prove that (1) implies (2), it suffices to note that the desired function M is Π_2^0 definable in G , and thus exists by \mathbf{ACA}_0 . Since the reachability matrix for G is Δ_1^0 definable in M , (2) implies (1) follows via an application of Theorem 1. ■

Imitating Corollary 2, one can easily formulate a version of part (2) of Theorem 4 for sequences of graphs which is provably equivalent to \mathbf{ACA}_0 . However, the use of Π_2^0 definability in the proof of Theorem 4 indicates a change in the recursion-theoretic content of the statement. This is explicitly demonstrated by the following result.

THEOREM 5. *There is a recursive graph G such that the function M of Theorem 4 is Π_2^0 complete.*

Proof. By the proof of Theorem 4, the counting function for any recursive graph is Π_2^0 definable. Thus it suffices to define a recursive graph G such that some Π_2^0 complete set is 1-reducible to M . Note that $\text{INF} = \{x : W_x \text{ is infinite}\}$ is Π_2^0 complete [5]. Let the vertices of G be of the form b_n, e_n , and $v_{n,j,k}$ for all $j, k, n \in \omega$. If $\{n\}(j)$ halts at precisely stage k , the edges $(b_n, v_{n,j,k})$ and $(v_{n,j,k}, e_n)$ are included in G . Clearly, G is recursive. Let M be the counting function for G . Then for all $n \in \omega$, $n \in \text{INF}$ if and only if $M(b_n, e_n) = \infty$. Thus $\text{INF} \leq_1 M$ as desired. (In this proof, G can be viewed as undirected or directed.) ■

A graph is *integer weighted* if there is a function $w : E \rightarrow \mathbb{N}^+$ that assigns a positive integer to each edge of G . If G is directed, the edges (u, v) and (v, u) could conceivably have distinct weights. \mathbf{RCA}_0 can prove that if two vertices of an integer weighted graph are connected by a path, then there is

a path of minimal total weight connecting them. A *shortest path matrix* for a graph G is a function $M : V \times V \rightarrow \mathbb{N} \cup \{\infty\}$ such that $M(v, w) = \infty$ if G contains no path from v to w , and otherwise $M(v, w)$ is a code for a path of minimal total weight. Including the code for a minimal weight path in the reachability matrix does not require additional proof-theoretic strength.

THEOREM 6. (RCA₀) *The following are equivalent:*

- (1) **ACA₀**.
- (2) *If G is a (2-regular) (directed) integer weighted graph, then there is a shortest path matrix for G .*

PROOF. To prove that (1) implies (2), it suffices to note that a shortest path matrix is Π_1^0 definable in G , thus exists by **ACA₀**. Since the reachability matrix of G is Δ_1^0 definable in any shortest path matrix, (2) implies (1) follows via an application of Theorem 1. ■

The reader may wish to verify that Corollaries 2 and 3 can be modified by replacing *reachability matrix* by *shortest path matrix*. Also, Theorem 6 holds for rationally weighted graphs, that is, graphs with positive rationals assigned to the edges.

Topological sorting. In this section we consider two infinite versions of topological sorting of directed graphs. The following definition is non-traditional in that it allows the same label to be assigned to many vertices.

DEFINITION 7. (RCA₀) A function $s : V \rightarrow \mathbb{N}$ is a *weak topological sorting* of the directed graph $G = (V, E)$ if

- (1) the range of s is either \mathbb{N} or an initial segment of \mathbb{N} , and
- (2) if $(u, v) \in E$, then $s(u) < s(v)$.

To simplify the following development, we restrict our attention to graphs with a source node. The vertex v_0 is a *source node* for G if for every $u \in V$ such that $u \neq v_0$, there is a directed path from v_0 to u . If v_0 is a source node for G , we say G is *path bounded* if for every $u \neq v_0$, there is an integer k such that the length of every path from v_0 to u is at most k . Clearly, path bounded graphs are acyclic. In fact, path boundedness is a necessary and sufficient condition for the existence of a weak topological sorting. Part of this equivalence is provable in **RCA₀**, as demonstrated by the following theorem.

THEOREM 8. (RCA₀) *If G is a directed graph which has a source node and a weak topological sorting, then G is path bounded.*

PROOF. Suppose G has source node v_0 and a weak topological sorting $s : V \rightarrow \mathbb{N}$. Suppose by way of contradiction that G is not path bounded. Then there is a vertex u such that there are arbitrarily long paths from v_0

to u . Suppose $s(u) = k$. Let v_0, v_1, \dots, v_j, u be a directed path such that $j > k$. By the definition of a weak topological sorting, $0 \leq s(v_0) < s(v_1) < \dots < s(v_j) < s(u)$, so $s(u) > j > k$, a contradiction. ■

Proving that path boundedness is sufficient to guarantee the existence of a weak topological sorting requires **ACA**₀.

THEOREM 9. (RCA₀**)** *The following are equivalent:*

- (1) **ACA**₀.
- (2) *Every path bounded directed graph with a source node has a weak topological sorting.*

PROOF. To prove that (1) implies (2), assume **ACA**₀ and let G be a path bounded directed graph with source node v_0 . Define $s : V \rightarrow \mathbb{N}$ by setting $s(v)$ to the least integer k such that every path from v_0 to V is of length at most k . By **ACA**₀, the function s exists. Whenever (u, v) is an edge of G , $s(v) \geq s(u) + 1$, so s is the desired weak topological sorting.

To prove the converse, assume **RCA**₀ and (2). As in the proof of Theorem 1, it suffices to prove the existence of the range of an injection $f : \mathbb{N} \rightarrow \mathbb{N}$. Let G be the directed graph with vertices $\{r_i \mid i \in \mathbb{N}\}$ and $\{d_i \mid i \in \mathbb{N}\}$. For each $i \in \mathbb{N}$, include the directed edges (r_i, r_{i+1}) in G . Additionally, if $f(j) = i$, include the edges (r_i, d_j) and (d_j, r_{i+1}) . **RCA**₀ proves that G exists and is a path bounded directed graph. Because the domain of f is \mathbb{N} , r_0 is a source node for G . Applying (2), let $s : V \rightarrow \mathbb{N}$ be a weak topological sorting of G . Since s maps V onto \mathbb{N} , for each i we have $s(r_{i+1}) = s(r_i) + 2$ if and only if i is in the range of f . Since the range of f is Δ_1^0 definable in s , **RCA**₀ proves its existence, completing the proof. ■

Note that one can still prove Theorems 8 and 9 if the first clause is omitted from Definition 7. To do this, the immediately preceding argument must be modified by inserting a chain of j nodes between r_i and r_{i+1} whenever $f(j) = i$.

The proof of Theorem 9 can be adapted to provide recursion-theoretic consequences.

COROLLARY 10. *Each path bounded directed recursive graph with a source node has a weak topological sorting recursive in $0'$. Also, there is a path bounded directed recursive graph G with a source node such that $0'$ is recursive in any weak topological sorting of G .*

PROOF. To prove the first statement, it suffices to note that the weak topological sorting constructed in the first part of the proof of Theorem 9 is Π_2^0 and Σ_2^0 definable. By Post's Theorem (see [5]), the result follows. To construct the recursive graph for the second statement, imitate the construction

in the second part of the proof of Theorem 9, letting f be some recursive injection with $0'$ recursive in its range. ■

Traditionally, topological sortings of graphs do not assign the same label to many nodes. Consequently, we will now consider the following definition.

DEFINITION 11. (\mathbf{RCA}_0) A function is a *topological sorting* if it is an injective weak topological sorting.

Not all graphs that have weak topological sortings have topological sortings. We say that a directed graph G with source node v_0 is *path sparse* if for every vertex $u \neq v_0$, there are at most finitely many directed paths from v_0 to u . Path sparseness plays the same role for topological sorting as path boundedness played for weak topological sorting, as reflected by the next two theorems.

THEOREM 12. (\mathbf{RCA}_0) *If G is a directed graph which has a source node and a topological sorting, then G is path sparse.*

Proof. Similar to the proof of Theorem 8. ■

THEOREM 13. (\mathbf{RCA}_0) *The following are equivalent:*

- (1) \mathbf{ACA}_0 .
- (2) *Every path sparse directed graph with a source node has a topological sorting.*

Proof. To prove that (1) implies (2), assume \mathbf{ACA}_0 and let $G = (V, E)$ be a path sparse graph with source node v_0 . Fix an enumeration of V in which each node occurs infinitely often. Define $s^{-1} : \mathbb{N} \rightarrow V$ and a sequence of integer markers $\langle j_n \mid n \in \mathbb{N} \rangle$ as follows. Let $s^{-1}(0) = v_0$ and $j_0 = 0$. Suppose that $s^{-1}(n)$ and j_n have been defined. Let S be the set of all vertices v such that v is not in $\{s^{-1}(k) \mid k \leq n\}$ and every vertex other than v lying on any path leading from v_0 to v is in $\{s^{-1}(k) \mid k \leq n\}$. Let $s^{-1}(n+1)$ be the vertex in S appearing first after v_{j_n} in the enumeration. Let j_{n+1} be the least integer greater than j_n such that $v_{j_{n+1}} = s^{-1}(n+1)$. \mathbf{ACA}_0 suffices to prove that s^{-1} exists. Since G is path sparse, use of the j_n s insures that s^{-1} is a bijection between \mathbb{N} and V . A straightforward induction argument shows that the inverse function $s : V \rightarrow \mathbb{N}$ is the desired topological sorting.

To prove that (2) implies (1), imitate the reversal of Theorem 9. ■

Recursion-theoretic analysis of the proof of Theorem 13 yields a version of Corollary 10 with weak topological sortings replaced by topological sortings.

Our results concerning topological sortings can easily be extended to sequences of graphs. The reader may wish to verify the following results.

COROLLARY 14. (\mathbf{RCA}_0) *The following are equivalent:*

(1) \mathbf{ACA}_0 .

(2) *For any sequence $\langle G_i \mid i \in \mathbb{N} \rangle$ of directed graphs with source nodes, there is a function $g : \mathbb{N} \rightarrow \{0, 1\}$ such that $g(i) = 1$ if and only if G_i has a (weak) topological sorting.*

One can construct a recursive sequence of graphs such that the deciding function of Corollary 14 for the existence of (weak) topological sortings is Π_3^0 complete.

Spanning trees. In this section we consider spanning trees. A *spanning tree* for a graph G is an acyclic connected subgraph of G that contains every vertex of G . For Theorem 15 through Theorem 19, we will restrict our attention to undirected graphs. A proof that every connected recursive graph has a recursive spanning tree can be derived from the following theorem.

THEOREM 15. (\mathbf{RCA}_0) *Every connected graph has a spanning tree.*

Proof. We work in \mathbf{RCA}_0 . Let G be a connected graph and let $\langle e_i \mid i \in \mathbb{N} \rangle$ be an enumeration of the edges of G . Define $h : \mathbb{N} \rightarrow \{0, 1\}$ by letting $h(0) = 1$ and setting $h(n) = 1$ if and only if the subgraph with edges $\{e_n\} \cup \{e_i \mid i < n \wedge h(i) = 1\}$ does not contain a cycle. \mathbf{RCA}_0 proves that h exists and is the characteristic function for the edge set of a spanning tree for G . ■

Every spanning tree is minimal in the sense that deletion of one edge yields a disconnected graph. Minimality is more subtle for edge weighted graphs. A graph is *rationally weighted* if there is a function $w : E \rightarrow \mathbb{Q}^+$ that assigns a positive rational to each edge. We say that a rationally weighted graph G has a *minimal spanning tree* if there is a spanning tree for G whose edge weights sum to a *finite* value which is no greater than the sum of the edge weights for any other spanning tree of G . The next theorem shows that connectedness of G does not guarantee the existence of a minimal spanning tree.

THEOREM 16. *There is a rationally weighted connected recursive graph with no minimal spanning tree.*

Proof. Let G be the ladder graph with vertices labeled l_i and r_i and edges (l_i, l_{i+1}) , (r_i, r_{i+1}) , and (l_i, r_i) for each $i \in \omega$. Define a weight function, w , for the edges by setting $w(l_i, l_{i+1}) = w(r_i, r_{i+1}) = 2^{-i}$ and $w(l_i, r_i) = 5 \cdot 2^{-i}$. Suppose T is a spanning tree for G . Then there is a lowest rung of the form (l_j, r_j) in T . Consider the graph T' obtained by deleting (l_j, r_j) from T and adding the edges (l_{j+1}, r_{j+1}) , (l_j, l_{j+1}) , and (r_j, r_{j+1}) . (One of these edges may be omitted if the resulting graph is not acyclic.) Note that T' spans G and $w(T') \leq w(T) - 2^{-(j+1)} < w(T)$, so T is not minimal. ■

We can formulate a sufficient condition for the existence of a minimal spanning tree using the following terminology. A graph G is *finitely connected* if it is connected and each pair of vertices is connected by only finitely many distinct arcs.

THEOREM 17. (\mathbf{RCA}_0) *The following are equivalent:*

- (1) \mathbf{ACA}_0 .
- (2) *Every finitely connected rationally weighted graph of bounded total weight has a minimal spanning tree.*

Proof. To prove that (1) implies (2), assume \mathbf{ACA}_0 , and let G be a finitely connected graph of bounded total weight. For any finite set \vec{e} of edges and any integer k , let $G(\vec{e}, k)$ denote the subgraph of G whose edges consist of \vec{e} together with those edges whose codes are greater than k . Fix an enumeration of all pairs of vertices from G . Build the tree T as follows. Each node of T is an arc between the pair of vertices corresponding to its level. If σ is a sequence of such nodes, then $\sigma \in T$ if and only if

$$\forall \vec{e} \forall k [(G(\vec{e}, k) \text{ spans } G) \rightarrow w(G(\vec{e}, k)) > w(\bigcup \sigma)].$$

Since G is finitely connected, T is finitely branching. Also, for the sequence of pairs of vertices of a given length l , there are only finitely many ways to assign arcs to the vertex pairs. Consequently, there is an assignment σ of length l which minimizes $w(\bigcup \sigma)$. Since every spanning graph must contain an arc between each pair of vertices, if $G(\vec{e}, k)$ spans G , then $w(G(\vec{e}, k)) > w(\bigcup \sigma)$. Thus, T contains a sequence of length l for arbitrary l . Summarizing, T is infinite and finitely branching. Applying König's Lemma (which is a consequence of \mathbf{ACA}_0) yields an infinite path through T . Let S denote the union of the path through T . Since an arc between each pair of vertices is included in S at the level of T associated with that pair, S is a spanning tree.

Suppose, by way of contradiction, that S is not minimal. Then there is a spanning tree S' of G and an $\varepsilon > 0$ such that $w(S') + \varepsilon < w(S)$. Choose k so large that ε is greater than the sum of the weights of all edges of G whose vertices have codes larger than k . Let \vec{e} be the collection of all edges of S' with codes less than or equal to k . Then $G(\vec{e}, k)$ spans G and $w(G(\vec{e}, k)) < w(S') + \varepsilon < w(S)$. Thus, for some initial segment σ of the path defining S , $w(G(\vec{e}, k)) < w(\bigcup \sigma)$, contradicting the construction of T . Thus, S is a minimal spanning tree.

To prove that (2) implies (1), we will use (2) to prove the existence of the range of an injection. Assume \mathbf{RCA}_0 and (2), and let $f : \mathbb{N} \rightarrow \mathbb{N}$ be an injection. Define the graph G as follows. The vertices of G will be $l_{i,j}$ and $r_{i,j}$ for each $i, j \in \mathbb{N}$. For each $i, j \in \mathbb{N}$, add the edges $(l_{i,j}, l_{i,j+1})$ and $(r_{i,j}, r_{i,j+1})$ to G , assigning the weights $w(l_{i,j}, l_{i,j+1}) = w(r_{i,j}, r_{i,j+1}) =$

2^{-i-j} . For each $i \in \mathbb{N}$, add the edges $(l_{i,0}, r_{i,0})$ and $(r_{i,0}, l_{i+1,0})$, assigning the weights $w(l_{i,0}, r_{i,0}) = w(r_{i,0}, l_{i+1,0}) = 2^{-i}$. Finally, whenever $f(s) = t$, add the edge $(l_{t,s+1}, r_{t,s+1})$ to G , with $w(l_{t,s+1}, r_{t,s+1}) = 2^{-t-1}$. Note that $w(G) \leq 13$, so G is bounded. Additionally, \mathbf{RCA}_0 can prove that G is finitely connected. By (2), G has a minimal spanning tree. For each $t \in \mathbb{N}$, t is in the range of f if and only if $(l_{t,0}, r_{t,0})$ is not in the minimal spanning tree. By recursive comprehension, the range of f exists. ■

In the terminology of recursive graph theory, a graph is *highly recursive* if there is a recursive function $f : V \rightarrow \mathbb{N}$ mapping the vertices of the graph into the integers such that $f(v)$ is an upper bound on the codes of all vertices adjacent to v . The following corollary is a recursion-theoretic recasting of the preceding result.

COROLLARY 18. *There is a bounded finitely connected highly recursive graph G such that $0'$ is recursive in every minimal spanning tree of G .*

PROOF. Imitate the reversal in the proof of Theorem 17, letting f be a recursive function such that $0'$ is recursive in the range of f . ■

If G is a weighted finite graph in which no two edges have the same weight, then G has a unique minimal spanning tree. Given the existence of a minimal spanning tree, \mathbf{RCA}_0 suffices to prove uniqueness in the infinite analog of this case, as shown below.

THEOREM 19. (\mathbf{RCA}_0) *Suppose that G is a rationally weighted graph in which no two edges have the same weight. If G has a minimal spanning tree, then G has a unique minimal spanning tree.*

PROOF. Suppose G is as hypothesized. By way of contradiction, suppose T_1 and T_2 are distinct minimal spanning trees. The recursive comprehension axiom suffices to prove the existence of the set S of those edges which lie in exactly one of T_1 and T_2 . Since T_1 and T_2 are distinct, S is non-empty. Furthermore, S must contain an edge e of maximum weight.

Without loss of generality, we may suppose that $e \in T_1$. Let T'_1 denote $T_1 - \{e\}$. \mathbf{RCA}_0 suffices to prove that T'_1 has exactly two components and that there is at least one edge $e' \in T_2$ such that $e' \notin T_1$ and one vertex of e' lies in each component. By the choice of e , $w(e') < w(e)$. Consequently, $w(T'_1 \cup \{e'\}) < w(T_1)$. Since $T'_1 \cup \{e'\}$ spans G , this contradicts the minimality of T_1 . Thus, G has a unique minimal spanning tree. ■

The condition in Theorem 17 is sufficient to prove the existence of minimal spanning trees, but not necessary. The next theorem shows that the existence of a choice function for graphs with minimal spanning trees is provable in $\mathbf{II}_1^1\text{-CA}_0$. For the next two theorems, we will consider both undirected and directed graphs. We define a *spanning tree for a directed*

graph as an acyclic directed spanning graph with a source node. Thus, for any vertex v in a directed graph G , any spanning tree for G must contain a directed path from the source node to v .

THEOREM 20. (Π_1^1 -CA₀) *If $\langle G_i \mid i \in \mathbb{N} \rangle$ is a sequence of rationally weighted (directed) graphs, then there is a function $f : \mathbb{N} \rightarrow \{0, 1\}$ such that $f(i) = 1$ if and only if G_i has a minimal spanning tree.*

PROOF. First we will prove the statement for directed graphs. Assume Π_1^1 -CA₀ and let $\langle G_i \mid i \in \mathbb{N} \rangle$ be a sequence of rationally weighted directed graphs. We will show that there is a Σ_1^1 formula with free variable n which is equivalent to “ G_n has a minimal spanning tree.” Since the naïve formalization of this statement is Σ_2^1 , some effort is necessary.

We will need to refer to subgraphs of the following form. If S is a subgraph of G_n , \vec{e} is a finite set of edges of G_n , and k is an integer, let $S(\vec{e}, k)$ denote the subgraph of G_n whose edges consist of \vec{e} together with those edges of S whose codes are greater than k . Let $\psi(T, n)$ be a formula stating that there exists an S such that

- S is a spanning tree of G_n of bounded total weight,
- T is a spanning tree of G_n , and
- for every \vec{e} and k , if $S(\vec{e}, k)$ has a source node and spans G_n , then $w(S(\vec{e}, k)) \geq w(T)$.

Note that $\psi(T, n)$ can be formalized as a Σ_1^1 statement with the sequence $\langle G_i \mid i \in \mathbb{N} \rangle$ as a parameter.

Now we show that $\psi(T, n)$ holds if and only if T is a *minimal* spanning tree of G_n . First, if T is a minimal spanning tree for G_n , setting $S = T$ makes all the clauses of the formula true, so $\psi(T, n)$ holds. Now suppose that T is a spanning tree of G_n , but not minimal. Then there is a spanning tree \widehat{T} of G_n such that $w(\widehat{T}) + \varepsilon = w(T)$ for some $\varepsilon > 0$. Let S be any spanning tree of G of bounded weight. Since the weight of S is bounded, there is a k such that $w(S(\emptyset, k)) < \varepsilon$. Let \vec{e}_1 be the collection of all edges of \widehat{T} included in paths from the source node of \widehat{T} to vertices which are end nodes of edges of S that have codes less than or equal to k . Let \vec{e}_2 be the collection of edges of \widehat{T} that lie on the path from the source node of \widehat{T} to the source node of S . Let $\vec{e} = \vec{e}_1 \cup \vec{e}_2$, and note that \vec{e} is finite. Also, note that $S(\vec{e}, k)$ is a spanning subgraph of G_n with a source node, and

$$w(S(\vec{e}, k)) < w(\widehat{T}) + w(S(\emptyset, k)) < w(\widehat{T}) + \varepsilon = w(T).$$

Thus, if T is a spanning tree which is not minimal, then $\neg\psi(T, n)$ holds. Combining this with the second clause of $\psi(T, n)$, we have shown that if T is not a minimal spanning tree, then $\neg\psi(T, n)$ holds. Summarizing, T is a minimal spanning tree of G_n if and only if $\psi(T, n)$.

To complete the proof for the directed case, note that $\exists T\psi(T, n)$ is a Σ_1^1 formula asserting that G_n has a minimal spanning tree. By $\Pi_1^1\text{-CA}_0$, there is a function $f : \mathbb{N} \rightarrow \{0, 1\}$ such that $f(n) = 1$ if and only if $\exists T\psi(T, n)$.

By using the technique of replacing undirected edges by two directed edges, the preceding argument can be extended to the undirected case. ■

A partial converse to Theorem 20 can be proved using the following lemma.

LEMMA 21. (\mathbf{RCA}_0) *Let $\langle T_n \rangle_{n \in \mathbb{N}}$ be a sequence of trees. There is a sequence $\langle T'_n \rangle_{n \in \mathbb{N}}$ of trees such that for each n , T_n is well founded if and only if T'_n is well founded, and for every m , T'_n contains a sequence of length m .*

Proof. Let p_m denote the m th prime. Let T be the tree consisting of sequences of the form $\langle p_m, p_m^2, \dots, p_m^j \rangle$ where $j \leq m$. Given a tree T_n , let T'_n be the disjoint union of T_n and T . ■

Now we can state and prove the partial reversal for Theorem 20. In order to emphasize the resulting equivalence result, we have included the restriction of Theorem 20 to directed graphs in the following statement.

THEOREM 22. (\mathbf{RCA}_0) *The following are equivalent:*

- (1) $\Pi_1^1\text{-CA}_0$.
- (2) *If $\langle G_n \rangle_{n \in \mathbb{N}}$ is a sequence of rationally weighted directed graphs, then there is a function $f : \mathbb{N} \rightarrow \{0, 1\}$ such that for all n , $f(n) = 1$ if and only if G_n has a minimal spanning tree.*

Proof. The proof of (1) implies (2) is as in Theorem 20.

To prove that (2) implies (1), it suffices to use (2) to select the well founded trees from a sequence of trees. By Lemma 21, we may assume that each tree contains arbitrarily long sequences. Given T_n , construct the directed rationally weighted graph G_n as follows. Let $\langle v_n \rangle_{n \in \mathbb{N}}$ be an enumeration of the sequences in T , with $v_0 = \langle \rangle$. These will serve as the vertices of G_n . Edges are added to the graph in two ways. First, if v_k extends v_j and $\text{lh}(v_k) = \text{lh}(v_j) + 1$, add the (upward) edge (v_j, v_k) to G_n , assigning the weight

$$w(v_j, v_k) = 2^{-\text{lh}(v_k)} + 2^{-k}.$$

Secondly, if $\text{lh}(v_k) = \text{lh}(v_j) + 1$, but v_j is not an initial segment of v_k , add the (downward) edge (v_k, v_j) to G_n , and set $w(v_k, v_j) = 2^{-j}$. The sequence $\langle G_n \rangle$ is Δ_1^0 definable in $\langle T_n \rangle$, so \mathbf{RCA}_0 proves that it exists.

By (2), there is a function f such that $f(n) = 1$ if and only if G_n has a minimal spanning tree. We will complete the proof by showing that G_n has a minimal spanning tree if and only if T_n is not well founded.

First, suppose that T_n is not well founded. \mathbf{RCA}_0 can prove that for every spanning tree S of G_n ,

- for each vertex v , S contains an edge of the form (u, v) , and
- for each m , S contains an edge of the form (u, v) where $\text{lh}(u) = m$ and $\text{lh}(v) = m + 1$.

Consequently, every spanning tree of G_n has a total weight of at least 2. Let $v_{k_0}, v_{k_1}, v_{k_2}, \dots$ be an infinite path through T_n where for each i , $\text{lh}(v_{k_i}) = i$. Let S be the subgraph of G_n containing upward edges of the form $(v_{k_i}, v_{k_{i+1}})$ for all i together with all downward edges of the form (v_{k_i}, v_j) where $\text{lh}(v_j) + 1 = \text{lh}(v_{k_i})$ and v_j is not an initial segment of v_{k_i} . \mathbf{RCA}_0 can prove that S spans G_n and the total weight of S is exactly 2. Thus, S is a minimal spanning tree for T_n .

Now suppose that T_n is well founded and S is a spanning tree of G_n . We show that S is not minimal. If for each j , S contains exactly one upward edge (u, v) such that $\text{lh}(v) = j$, then these edges must form a directed arc, encoding an infinite path through T_n and contradicting the assumption that T is well founded. Thus, there is a least j such that S contains at least two upward edges (u, v) and (\tilde{u}, \tilde{v}) with $\text{lh}(v) = \text{lh}(\tilde{v}) = j$. Since T_n contains arbitrarily long sequences, let $v_{k_0}, v_{k_1}, \dots, v_{k_{j+1}}$ be a path of length $j + 1$ through T_n . Let \widehat{S} be a graph consisting of

- edges of S of the form (u, v) where $\text{lh}(v) > j$,
- the upward edges $(v_{k_i}, v_{k_{i+1}})$ for $i \leq j$, and
- all downward edges of the form (v_{k_i}, u) for $i \leq j + 1$.

\mathbf{RCA}_0 can prove that \widehat{S} spans G_n and that $w(\widehat{S}) \leq w(S) - 2^{-j-1}$. Thus S is not minimal, completing the proof that T_n is well founded if and only if G_n has a minimal spanning tree. ■

The preceding argument proves the statement resulting from replacing “minimal spanning tree” with “spanning tree with weight exactly 2.” Furthermore, “exactly 2” can be replaced by any desired value. We conjecture that Theorem 22 also holds for undirected graphs, but will require a substantially different proof.

References

- [1] J. Hirst, *Combinatorics in subsystems of second order arithmetic*, Ph.D. Thesis, The Pennsylvania State University, 1987.
- [2] —, *Connected components of graphs and reverse mathematics*, Arch. Math. Logic 31 (1992), 183–192.
- [3] S. Simpson, *Which set existence axioms are needed to prove the Cauchy/Peano theorem for ordinary differential equations?*, J. Symbolic Logic 49 (1984), 783–802.

- [4] S. Simpson, *Subsystems of Z_2* , in: Proof Theory, G. Takeuti (ed.), North-Holland, Amsterdam, New York, 1985, 434–448.
- [5] R. Soare, *Recursively Enumerable Sets and Degrees*, Springer, Berlin, 1987.

Department of Computer Science
Boston College
Chestnut Hill, Massachusetts 02167
U.S.A.
E-mail: clote@cs.bc.edu

Department of Mathematical Sciences
Appalachian State University
Boone, North Carolina 28608
U.S.A.
E-mail: jlh@math.appstate.edu

*Received 19 January 1997;
in revised form 23 January 1998*