The Proof of $P = NP$.
by
Cynthia Ann Harlan Krieger
Dr. Lee K. Jones
University of Massachusetts, Lowell, MA 01854

1 Theorem: P=NP.

The problem posed by the Hamiltonian cycle question is to determine whether or not there exists within $G = (V, E)$ a Hamiltonian cycle, which is a simple cycle that contains each vertex in V, in polynomial running time or show that it cannot be done.

As stated on page 979 of Cormen et. al. (2001), one solution is to check all possible permutations of the vertices and then check to see if any of these permutations are a Hamiltonian path. However this naive approach has an exponential running time. We will not employ this technique here. An algorithm is given for solving the Hamiltonian cycle problem for any undirected graph $G = (V, E)$ with a worst-case running time bounded by $O(|V|^4)$. This development of a polynomial time algorithm that decides the Hamiltonian cycle problem, which is known to belong to the $NP-$complete class of languages, will prove that $P = NP$.

2 Proof:

In essence the goal of the NP, P problem has been to find the criteria by which to reduce the search space, or to rephrase it, to find the criteria by which the problem will always be a tautology. If we know the criteria for which the problem will always be a tautology then we always can always determine the answer, we just compare the two. My previous instructor Dr. John Sieg said that it would take some additional structural knowledge requirement of the NP-complete problem if we were ever going to be able to solve NP-complete problems in polynomial time. The criteria for the problem to always be a tautology is the infimum of the problem's set of all possible answers.

This is easily computed when translated into a distance, or metric definable problem as in the case of the existence of a Hamiltonian cycle of length $|V|$ within an undirected graph $G = (V, E)$, which we recast as seeing if the infimum of all possible traversed path lengths is equal to $|V|$, or equivalently the maximum of all simple paths traveled is of length $|V|$.

We can view this problem in terms of the Conjunctive Normal Formula structure, or cnf-formula, which comprises several clauses connected with "and" operations, where each clause is composed of a combination of literals connected with "or" operations. If the graph contains a negative weight cycle this analogous to the clause simultaneously containing both a literal and that respective literal's complement. When this happens the clause is always false indicating the failure of the entire statement, due to each clause being connected with the "and" operation. Similarly, it is impossible to find a "shortest-path" in a graph which contains a negative weight-cycle. A graph which contains a positive weight cycle, is analogous to a clause that contains a repeat of the same variable. We can eliminate that repetition without changing the overall true value of the statement.

2.1 Lemma:

We define the infimum of a vertex $v_i \in V$, denoted $\inf(v_i)$, to be the length of the maximum of all the simple paths able to be traversed from that node. We define the set containing the union of all these infimum and denote it $V_{\text{inf}} = \cup\{\inf(v_i)\}$, for all $v_i \in V$. We define the infimum of the undirected graph $G = (V, E)$, $\inf(G) = \inf(V_{\text{inf}})$.

There exists a Hamiltonian cycle within $G = (V, E)$, if and only if for one vertex $v_i \in V$, $|V| = \inf(v_i)$. It is obvious, from the definition of the infimum, that when a Hamiltonian cycle is in a graph it is true for the undirected graph $G = (V, E)$, that $\inf(G) = |V|$. This is exactly equal to saying the infimum of the set of all possible path lengths generated by that undirected graph $G = (V, E)$ is equal to $|V|$.

2.2 Forward Proof of Lemma:

If there exists a Hamiltonian cycle in an undirected graph $G = (V, E)$ then there exists a path $v_i \rightsquigarrow v_i$ such that it's infimum (or equivalently the maximum simple path length traversable from that node) is equal to $|V|$. Since the existence of a Hamiltonian cycle, means that there exists a path $v_i \rightsquigarrow v_i$, now all that I need to show is that $|V|$ equals the infimum or is the maximal simple path traversed from that node. Assume for the sake of contradiction that there exists a Hamiltonian cycle in graph $G = (V, E)$, and we consider all paths of the form $v_i \rightsquigarrow v_i$, however there is no path for which the infimum of those vertices, or the maximal simple path of length, equals $|V|$.

A path can fail to be a maximal simple path in two ways:

1) one it can not be a maximal path;

2) it could not be a simple path.

A simple path of size less than $|V|$ cannot possibly incorporate all of the vertices and hence this path isn't Hamiltonian. A path whose length is larger than $|V|$ cannot be simple since there are only $|V|$ nodes, if the path is longer than $|V|$ some nodes are repeated which again makes this path not a Hamiltonian cycle. This implies that a Hamiltonian cycle on the graph above, if it exists must have a maximum shortest path length $|V|$, and there are no other simple paths bigger than this value that are also simple paths, hence it is the infimum of the set of possible path lengths. But this is a contradiction.

2.3 Backward Proof of Lemma:

If we are given a graph where there exists vertex $v_i \in V$ such that the infimum of the path lengths traversable from that vertex is equal to $|V|$, (which is equivalent to saying it is true that the path $v_i \rightsquigarrow v_i$ is a maximum simple path whose length $|V|$), then this implies $G = (V, E)$ has a Hamiltonian cycle. If we have a vertex $v_i \in V$, such that the infimum of the set of all possible path lengths traversable from that vertex v_i is equal to $|V|$, then by definition the path given from $v_i \rightsquigarrow v_i$, meets all the criteria necessary for a Hamiltonian cycle to be in the undirected graph $G = (V, E)$. An infimum of the possible path lengths is a maximum simple path. It is a cycle, it traverses all the nodes, and it is simple hence it is a Hamiltonian cycle by definition.■

3 Proposed Algorithm:

This implies the technique that if I can find any one of maximal simple paths traversable from any node equal to the length $|V|$ I can determine that there exists a Hamiltonian cycle in the graph. However the Hamiltonian cycle, is by symmetry of ideas perceived as a maximal shortest simple paths traversable in a graph of length $|V|$ where there are at most no moves left; or can be perceived, with all edge weights equal to one when traversing in one direction and equal to negative one when traversing in another direction, as a shortest simple path having length 0 with at most $|V|$ more move left (to maintain equilibrium). These are two equivalent conditions upon the original graph $G = (V, E)$. We choose to derive two techniques based on the later view of the problem. So in summary we are searching for a shortest simple path of length 0 from $v_i \rightsquigarrow v_i$ reachable in at most V steps.

Given the undirected graph $G = (V, E)$, we begin our algorithm by constructing the adjacency matrix for the graph. We then modify the matrix as follows. Instead of the zeros down the diagonal we place $|V|$ down the diagonal. Then for all the values below the main diagonal, we replace the ones with negative ones, leaving the zeros below the diagonal untouched.

For the purposes of our proof, we decide the Hamiltonian cycle problem by using a modified shortest path algorithm (whose matrix multiplication is defined below) where we defined, through the adjacency matrix modifications, all the edges in E to have weight one, zero otherwise. Naturally, when using a shortest path algorithm the question of a negative weight cycle appears, by our addition of a weight of $|V|$ on each vertex the instance of a negative weight cycle happening is impossible.

We use a modified shortest path algorithm because a shortest path cannot contain a positive-weight cycle either, since removing the cycle results in a path from the same source to the same destination with a lower value, making the new path without the positive-weight cycle the simplest path violating the assumption that the path that contained the positive-weight cycle was a simple path with which to begin. Since this leaves only zero-weight cycles, we claim can remove these without affecting the weight of the path, so we do so without loss of generality. This implies that the simple path, having no positive or negative weight cycles, we are looking for can be found using the shortest path algorithm on the modified matrices above.

We modify the theory behind the All-Pairs Shortest Paths given in Chapter 25 of Cormen, et. al. (2001) and define the products $L^{(n-1)}$, where $L^{(n-1)} = W^{(n-1)} \cdot L^{(0)}$, and W_i is the adjacency representation of the graphs $G = (V, E)$ as modified above, and $[L^{(0)}]_{ij} = \{1 \text{ if } i = j, \text{ and } \infty \text{ if } i \neq j\}$. We now define the matrix product we will be using in this paper, denoted $C = A \cdot B$, to be the matrix product returned by $[C]_{ij} = abs(\min([C]_{ij}, [A]_{ik} + [B]_{kj}))$ This resulting "product", (as defined above), which we will call $L_i^{(n-1)}$, contains the shortest simple path for every pair of vertices i, and j that can be reached within at most $n-1$ moves that will allow cycles. However, by our weighting each node with a weight of $|V|$, if it returns to that node prematurely it will be penalized.

If we are looking for the maximum simple path from $v_i \rightsquigarrow v_i$ it makes sense to look at the diagonal values of the $L^{(n)}$ computations. If there is a shortest path $v_i \rightsquigarrow v_i$, reachable in $|V|$ moves it will be contained within these values. Therefore we compute the value $L^{(V)}$ using a matrix multiplication modified version of the Faster-All-Pairs-Shortest-Paths algorithm on page 627 of Cormen, et. al (2001). What we want is to see is if this computation results in a simple path able to be traversed in at most $|V|$ steps, from all vertices that is equal to 0.

By our two assumptions, none of the diagonal elements of $L^{(n)}$ are smaller than 0 because the implementation of the modified all-pairs shortest-paths algorithm does not return paths containing cycles, and that the definition of the problem that only one edge exist between all possible pairs of vertices. If any of the values on the diagonal of $L^{(n)}$ are larger than zero then the original undirected graph $G = (V, E)$ does not contain a Hamiltonian cycle, by Lemma above. Hence we conclude that if there exists a Hamiltonian path in the original graph structure then all the values on the diagonal of $L^{(n)}$ will be equal 0.

3.1 Worst-case Running Time Analysis:

All that remains to be shown is that this algorithm implements in polynomial time. If we use the adjacency representation, this construction takes time

$O(V^3)$. We assume $O(V^2)$ time to make the modifications to the adjacency matrix representation as specified in the algorithm. We assume a running time of the Faster-All-Pairs-Shortest-Paths(L) of $O(V^3 \lg V)$. We assume linear time to compare the elements on the diagonal of $L^{(n)}$ to 0 to decide whether or not there exists a Hamiltonian cycle found in graph $G = (V, E)$. If we sum up all of these times, we obtain a worst-case running time of $O(V^3 + V^2 + V^3 \lg V + V)$. This algorithm is bounded above by a worst-case running time of $O(V^4) \in P$.

4 Conclusions:

In essence the realization that we want to avoid all extraneous searches which contain cycles, immediately test the problem to see if it contains a negative weight cycle in which case we reject, and determine the set of all possible minimally sufficient answers and then take the maximum of that value and compare it to our problem at hand reduces the search space to something that can be computed in polynomial time. If our answer is the infimum of all possible answers then there cannot be a possible case where our criteria fails, and any value of the statement is larger than our value making the statement satisfactory. If the infimum of the solution set is higher than our chosen criteria then our solution could not be a tautology, because there are cases where our assignment of the variables will return a false solution.

So I argue that if the value we are looking for is exactly the infimum of the set of possible answers of the statement, then the answer to the problem is yes otherwise the answer to the problem is no. So now the question becomes to find the infimum of the set of answers and compare it to the value we are specifically looking for within the structure of the problem. One way to do so is to recast the nature of the problem into something with a metric which is readily, and quickly computable. For example, if we are looking to determine whether or not a Hamiltonian cycle of length $|V|$ exists within a graph, we merely look at the infimum of path lengths, or the maximum simple paths possible.

The modified adjacency matrix product $L^{(V)}$ contains the shortest simple path for every pair of vertices i, and j that can be reached within at most V moves. If there exists a Hamiltonian cycle in the graph we know that at all of the maximal simple paths generated from the above All-Source Shortest-Path constructed algorithm and listed on the diagonal of $L^{(n)}$ must be equal to 0, since this is just a rephrasing of the Hamiltonian problem where the path must equal $|V|$ when it has no moves left. If it is we are done and the answer is yes, if it is not we are still done and the answer is no. It is decided.

This works fast because the All-Pairs Shortest-Path algorithm examines only the possible permutations of vertices which do not contain cycles. Our modification does allow cycles, but it heavily penalizes them by putting the most weight on the vertex, a weight of equal to $|V|$. In short, we are vastly reducing the space of the exhaustive search done by the naive approach. We determine the set of minimal answers through the inherent property of the one step All-Pairs Shortest-Path approach which eliminates all the permutations of vertices that contain cycles.

Hence we have shown the existence of a polynomial time algorithm to determine the existence of a Hamiltonian cycle for the general undirected graph $G = (V, E)$. Since this problem is know to belong to the class of $NP-$complete languages, there exists for every language in the class of NP languages a polynomial time reduction to this problem. We would merely compose these two polynomial time functions, the reduction from the Hamiltonian problem to the NP language problem, and the solution provided to the Hamiltonian cycle problem contained in this paper to obtain a polynomial time bounded solution to every language in the NP class of languages. Hence we conclude $P = NP$.

5 References:

Cormen, T.H., Leiserson, C.E., Rivest, R. L., Stein, C. (2001), Introduction to Algorithms, (second edition), MIT Press, Cambridge, MA.