# Hamiltonian cycle

Vassos Hadzilacos

The ***Directed Hamiltonian Cycle*** problem, abbreviated DHC, is the following decision problem:

**Instance:** $\langle G \rangle$, where $G$ is a directed graph.

**Question:** Does $G$ have a simple cycle that visits every node? (A cycle $u_1, u_2, \ldots, u_k, u_1$ is ***simple*** if the nodes $u_1, \ldots, u_k$ are all distinct.)

A simple cycle that includes every node is called a ***Hamiltonian cycle***, and a graph that has such a cycle is called a ***Hamiltonian graph***. Figure 1 shows a Hamiltonian and a non-Hamiltonian graph.

Figure 1: A non-Hamiltonian graph (left) and a Hamiltonian graph (right)

**Theorem 10.3** DHC *is* ***NP****-complete.*

PROOF. It is straightforward to show that DHC $\in$ **NP**. Let $G = (V, E)$, and let $|V| = n$, $|E| = m$. The certificate is a sequence of nodes $u_1, u_2, \ldots, u_n$; this can be represented as a string of $O(m \log n)$ bits. The verifier checks that the nodes in the sequence are pairwise distinct, and that, for every $i \in [1..n-1]$, $(u_i, u_{i+1})$ is an edge of $G$, and that $(u_n, u_1)$ is also an edge of $G$. This can be done in polynomial time in $n$ and $m$.

We prove that DHC is **NP**-hard by showing that VERTEXCOVER $\leq_m^p$ DHC.

Given $\langle G, k \rangle$ where $G = (V, E)$ is an undirected graph and $k$ is an integer in $[1..|V|]$, we show how to construct, in polynomial time, a directed graph $G_D = (V_D, E_D)$ such that

$$G \text{ has a vertex cover of size } k \iff G_D \text{ has a Hamiltonian cycle.} \qquad (*)$$

To define the nodes and edges of $G_D$ we need some notation. We abbreviate the edge $\{u, v\}$ of $G$ as $uv$; since $G$ is undirected, $uv$ is exactly the same edge as $vu$. We list the edges of $G$ adjacent to node $u$ in some arbitrary order and denote them as $e_u^1, e_u^2, \ldots, e_u^{d_u}$, where $d_u$ is the degree of node $u$, i.e., the number of edges incident on $u$. The edge $uv$ is listed both among the edges adjacent to $u$ and also among the edges adjacent to $v$, so $uv$ is $e_u^i$ for some $i \in [1..d_u]$ as well as $e_v^j$ for some $j \in [1..d_v]$.

We now describe the nodes and edges of the directed graph $G_D$.

- $G_D$ has the following nodes:
  - $k$ nodes denoted $c_1, \ldots, c_k$, which we will call "cover" nodes, and

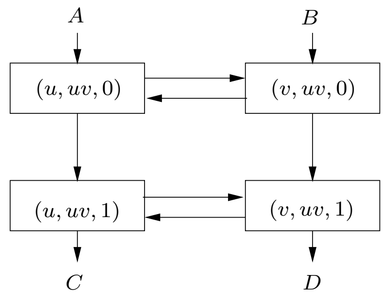


Figure 2: The four nodes of $G_D$ that correspond to the edge $uv$ of $G$

  - four nodes for every edge $uv$ of $G$, denoted $(u, uv, 0)$, $(u, uv, 1)$, $(v, uv, 0)$, and $(v, uv, 1)$. Getting a little ahead of ourselves, these four nodes will be connected as shown in Figure 2, with the edges coming from points $A$ and $B$ and going to points $C$ and $D$ to be explained shortly. Imagine the nodes of $G_D$ of the form $(u, -, -)$ being arranged vertically in a column in the order $(u, e_u^1, 0)$, $(u, e_u^1, 1)$, $(u, e_u^2, 0), (u, e_u^2, 1), \ldots,$ $(u, e_u^{d_u}, 0)$, $(u, e_u^{d_u}, 1)$.

- $G_D$ has the following edges:
  - For each $i \in [1..k]$ and each $u \in V$, the edge $\big(c_i, (u, e_u^1, 0)\big)$ — i.e., edges from each "cover" node $c_i$ to the first node of the column of $G_D$ nodes that corresponds to each node $u$ of $G$.
  - For each $i \in [1..k]$ and each $u \in V$, the edge $\big((u, e_u^{d_u}, 1), c_i\big)$ — i.e., edges from the last node of the column of $G_D$ nodes that corresponds to each node $u$ of $G$ to each "cover" node $c_i$.
  - For each $uv \in E$, the edges
    - $\big((u, uv, 0), (u, uv, 1)\big)$ and $\big((v, uv, 0), (v, uv, 1)\big)$ — the vertical edges shown in Figure 2;
    - $\big((u, uv, 0), (v, uv, 0)\big)$, $\big((u, uv, 1), (v, uv, 1)\big)$, $\big((v, uv, 0), (u, uv, 0)\big)$, $\big((v, uv, 1), (u, uv, 1)\big)$ — the horizontal edges shown in Figure 2.
  - For each $u \in V$ and $i \in [1..d_u - 1]$, the edge $\big((u, e_u^i, 1), (u, e_u^{i+1}, 0)\big)$ — the edges from $A$ and $B$ to $C$ and $D$, respectively, shown in Figure 2.

An example of the construction of $G_D$ from $G$ is shown in Figure 3. You may also find useful the step-by-step illustration of the construction in this example described here.

Let us first examine the time needed to construct $G_D$ from $G$. We have

$$|V_D| = 4m + k$$
$$|E_D| = 2kn + 6m + \sum_{u \in V} (d_u - 1) = 2kn + 6m + 2m - n = (2k - 1)n + 8m.$$

Without loss of generality we can assume that $k \leq n$: otherwise the given instance of Vertex Cover is obviously a no-instance and therefore we can map any such instance to a trivial no-instance of DHC. Therefore, $|V_D| = O(m + n)$ and $|E_D| = O(n^2 + m) = O(n^2)$. So the size of $G_D$ is polynomial in the size of $G$, and obviously it can be constructed from it in polynomial time.

It remains to prove $(*)$.

[Only If] Let $u_1, \ldots, u_k$ be a vertex cover of $G$. We will show that $G_D$ has a Hamiltonian cycle.

Consider the following path: Start at $c_1$, continue to $(u_1, e_{u_1}^1, 0)$ (the first node in the "column" of $G_D$ nodes that corresponds to the first node $u_1$ of the vertex cover of $G$), and then visit every node of the form $(u_1, -, -)$ in turn, following the "vertical" edges of that column. When the last node $(u_1, e_{u_1}^{du}, 1)$ of

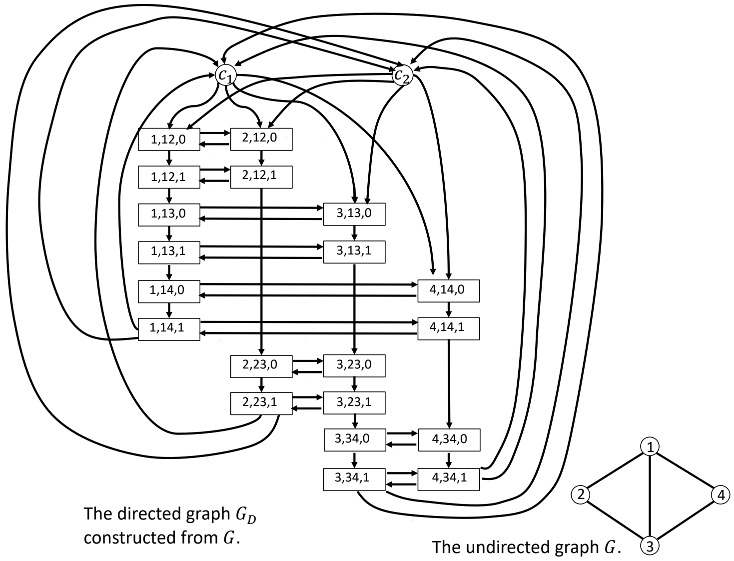


Figure 3: The directed graph $G_D$ obtained from the undirected graph $G$

that column is reached follow the edge to $c_2$, continue to $(u_2, e^1_{u_2}, 0)$ (the first node in the "column" of $G_D$ nodes that corresponds to the second node $u_2$ of the vertex cover of $G$), and then visit the nodes of the form $(u_2, -, -)$. After visiting these, follow the edge to $c_3$ and so on, until we have done the same with each node $u_i$, $i \in [1..k]$, in the vertex cover of $G$. From the last node of the column of nodes of the form $(u_k, -, -)$, return to $c_1$.

The path described above is a simple cycle, but it is not a Hamiltonian cycle because it misses the nodes of the form $(v, e^j_v, b)$ for all $v \neq u_i$, $i \in [1..k]$, $j \in [1..d_v]$, and $b \in \{0, 1\}$ — i.e., the nodes in the columns that do not correspond to nodes of $G$ in the vertex cover. Consider any such node, say $(v, e^j_v, b)$. Recall that $e^j_v$ is the edge $vu$ in $G$, for some node $u$; and since $v$ is not in the vertex cover of $G$, $u$ must be. So, $e^j_v = e^i_u$ for some $u$ in the vertex cover and $i \in [1..d_u]$. Thus, we can modify the above path to include the nodes $(v, e^j_v, b)$ by replacing the edge $(u, e^i_u, 0), (u, e^i_u, 1)$ by the path $(u, e^i_u, 0), (v, e^j_v, 0), (v, e^j_v, 1), (u, e^i_u, 1)$. (See Figure 2: instead of going directly down from $A$ to $D$, we take a detour to include the two nodes on the right).

By adjusting the path in this manner for all the nodes it misses, we obtain a Hamiltonian cycle of $G_D$.

[If] Suppose that $H$ is a Hamiltonian cycle of $G_D$. We will show that $G$ has a vertex cover of size $k$.

The cycle $H$ must pass through all the nodes $c_1, \ldots, c_k$ in some order. Without loss of generality, assume that it does so in this order (we can ensure this by re-indexing the nodes $c_1, \ldots, c_k$, if necessary). So, $H$ consists of $k$ segments, each starting at $c_i$ and ending in $c_{i \oplus 1}$, for $i \in [1..k]$, where $i \oplus 1 = (i \mod k) + 1$ (so the "next" integer after $k$ is 1):

$$H = c_1 \leadsto c_2 \leadsto c_3 \leadsto \cdots \leadsto c_k \leadsto c_1.$$

From the definition of $G_D$, the first node after $c_i$ on the $c_i \leadsto c_{i \oplus 1}$ segment of $C$ is $(u_i, e^1_{u_i}, 0)$, for some node $u_i$ of $G$. We will show that $u_1, u_2, \ldots, u_k$ form a vertex cover of $G$.

To see why, first refer to Figure 2. If $H$ enters this group of four nodes from $A$, it must exit from $C$: if it exits from $D$ it will miss one of the other two nodes of the group. Similarly, if $H$ enters this group of four nodes from $B$, it must exit from $D$. Therefore,

$$\text{every node on the } c_i \leadsto c_{i \oplus 1} \text{ segment of } H, \text{ except } c_i \text{ and } c_{i \oplus 1}, \text{ is of the form } (-, u_i v, -)$$

(recall that $u_i v$ is identical to $v u_i$). This implies that

$$\text{for every node } (-, e, -) \text{ on } H, \text{ the edge } e \text{ of } G \text{ is covered by one of the nodes } u_1, \ldots, u_k.$$

Since $H$ passes through every node of $G_D$, and for every edge $e$ of $G$ there are nodes $(-, e, -)$ in $G_D$, it follows that $u_1, \ldots, u_k$ is a vertex cover of $G$, as wanted. □

## The undirected Hamiltonian cycle problem

The undirected Hamiltonian cycle problem, UHC, is just like DHC, except that the graph $G$ is undirected. Note that a cycle in an undirected graph must have length at least three; that is, if $\{u, v\}$ is an edge of $G$, $u, v, u$ is not a cycle. (In contrast, a directed graph can have cycles of length 2.) Figure 4 shows two undirected graphs, one that has no Hamiltonian cycle and one that does.

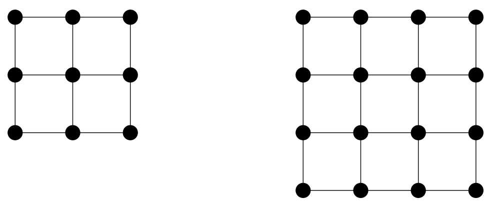

Figure 4: Undirected graphs without (left) and with (right) Hamiltonian cycle

**Theorem 10.4** UHC *is **NP**-complete.*

Proof Sketch. It is straightforward to show that UHC is in **NP**. To show that it is **NP**-hard, we sketch a polytime mapping reduction of DHC to UHC, leaving the detailed argument as an exercise.

Given a directed graph $G = (V, E)$ we construct an undirected graph $G' = (V', E')$ such that $G$ has a Hamiltonian cycle if and only $G'$ does. Intuitively, the idea is to create three nodes $u_1, u_2, u_3$ in $G'$ for each node $u$ of $G$. We add edges $\{u_1, u_2\}$ and $\{u_2, u_3\}$, and for every (directed) edge $(u, v)$ of $G$ we add the (undirected) edge $\{u_3, v_1\}$ in $G'$. This construction is illustrated in Figure 5.

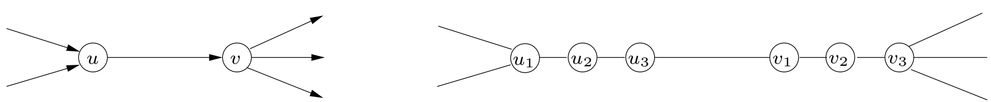


Figure 5: Illustration of reduction of DHC to UHC

More precisely, if $G = (V, E)$, we define $G' = (V', E')$ as follows:

$$V' = V \times \{1, 2, 3\}$$
$$E' = \big\{\{(u,1),(u,2)\}, \{(u,2),(u,3)\} \colon u \in V\big\} \cup \big\{\{(u,3),(v,1)\} \colon (u,v) \in E\big\}$$

It is obvious that $G'$ can be constructed in time polynomial in the size of $G$. We leave it as an exercise to prove that $G$ has a Hamiltonian cycle if and only if $G'$ does. The only-if direction is straightforward. The converse is a little more delicate. (Check that your proof does not apply if instead we had "split" each node $u$ of $G$ into two, rather than three, nodes in $G'$. Show, by means of a counterexample, that this simpler construction is ***not*** a correct reduction.) □