CPSC 320 2023S2: Tutorial 4 Solutions

1 Hamming Labelings of Graphs

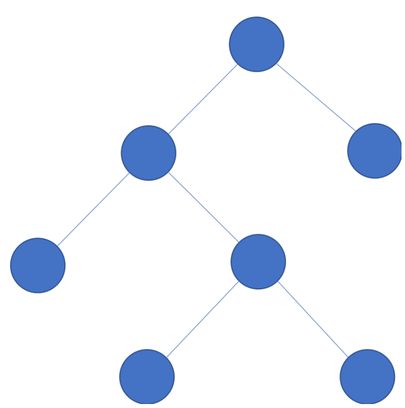
Let $G = (V, E)$ be an undirected, unweighted, connected graph with n nodes and m edges. Assume that G has no self-loops or multiple edges. We want to label the nodes in V with equi-length binary strings so that the length of the shortest path between any pair of nodes equals the the Hamming distance between the corresponding node labels. We call such a labeling a *Hamming labeling*.

Here, the Hamming distance between two length-k binary strings $x_1x_2...x_k$ and $y_1y_2...y_k$ is the number of positions $i, 1 \leq i \leq k$ where bits x_i and y_i differ. Given a Hamming labeling of a graph, we denote the label of node v by $l(v)$. Also, the length of a path between two nodes is the number of edges along the path, and we denote the length of the shortest path between nodes u and v by $d(u, v)$.

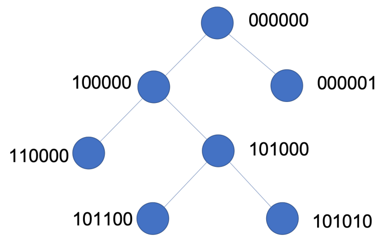
For example, consider the "square" graph where $V = \{1, 2, 3, 4\}$ and $E = \{(1, 2), (2, 3), (3, 4), (4, 1)\}$. The labeling $l(1) = 00$, $l(2) = 01$, $l(3) = 11$ and $l(4) = 10$ is a Hamming labeling.

However, the "triangle" graph with $V = \{1, 2, 3\}$ and $E = \{(1, 2), (2, 3), (3, 1)\}$ has no Hamming labeling. To see why, suppose to the contrary that the triangle graph does have a Hamming labeling $l()$. Without loss of generality (by reordering bits of the labeling, or flipping a given bit in all labels and/or renumbering the nodes) we can assume that $l(1) = 00...0$ (i.e., the string of all 0's), and that $l(2) = 100..0$ (i.e., the string consisting of a single 1 followed by all 0's). Also, $l(3)$ must have exactly one 1. We can consider two cases. The first is that $l(3)$ also starts with a 1. But then $l(3) = l(2)$, in which case $H(l(2), l(3)) = 0$ while $d(2, 3) = 1$, contradiction. Alternatively, $l(3)$ does not start with a 1. In this case, $H(l(2), l(3)) = 2$ while $d(2, 3) = 1$, again a contradiction. We conclude that the triangle graph does not have a Hamming labeling.

1. Provide a Hamming labeling for the following tree. See if you can find one with labels of length 6.



SOLUTION: Here's one possible labeling:



2. Suppose that graph G has a Hamming labeling. Show that G must be bipartite, i.e., has no odd-length cycle. (There's a hint at the top of the next page; you might want to play with some examples and try coming up with your own approach before you read the hint.)

   SOLUTION: Here's a nice proof suggested by TA Charles Hu that's simpler than the original proof in these solutions. The original proof is below.

   Suppose to the contrary that G is not bipartite. Let $v_1, v_2, ..., v_{2k+1}$ be an odd-length cycle in G, with labels $l(v_1), l(v_2), ..., l(v_{2k+1})$ in the Hamming labeling. The parity, i.e., number of 1-bits, in the labels of successive nodes on the cycle *must* alternate, since the label of successive nodes changes in exactly one position. So if the parity of $l(v_1)$ is even, then the parity of $l(v_2)$ is odd, and so on, so that the parity of $l(v_{2k+1})$ is even. But then the labeling is not a Hamming labeling since the two adjacent nodes v_1 and v_{2k+1} either have the same labeling or their labels differ by at least 2. The argument is similar when the parity of $l(v_1)$ is odd: in this case the parity of $l(v_{2k+1})$ is also odd and again the two adjacent nodes v_1 and v_{2k+1} either have the same labeling or their labels differ by at least 2. This contradicts the assumption that the labeling is a Hamming labeling.

And here's the original proof: Suppose to the contrary that G is not bipartite. Then G has an odd cycle. Choose the shortest such cycle, and let its length be $2k+1$, where $k \geq 1$. Pick some arbitrary node of the cycle; we'll call this node a. Without loss of generality we can assume that $l(a) = 00..0$. Let b and c be the two nodes of the cycle of distance exactly k from node a. Note that there is an edge between nodes b and c, so $d(b,c)$, which is the length of the shortest path from b to c, is 1.

We claim that the label $l(b)$ must differ by k bits from $l(a)$. To see why, suppose that to the contrary it differs by fewer bits, in which case there is a path $a, v_1, ...v_{l-1}, b$ from a to b of length $l < k$. We will construct an odd-length cycle in G of length less than $2k+1$, contradicting the fact that $2k+1$ is the length of the shortest odd-length cycle. Let $a = u_0, u_1, ..., u_k = b, u_{k+1}, ..., u_{2k}, u_0$ be the shortest cycle chosen initially. Then either $u_0, u_1, ..., u_k, v_{l-1}, ..., v_1, u_0$ or $v_0, v_1, ...v_{l-1}, u_k, ..., u_{2k}, u_0$ is a shorter odd-length cycle in G.

Without loss of generality, suppose that $l(b) = 11..100..0$, where the first k bits are 1's and the remaining bits are 0's. Also, the label $l(c)$ must differ by k bits from $l(a)$.

Suppose that there are j common positions where both $l(b)$ and $l(c)$ are 1. Then there are $k-j$ positions where $l(b) = 1$ and $l(c) = 0$, and another $k-j$ positions where $l(c) = 1$ and $l(b) = 0$. So $H(l(b), l(c)) = 2(k-j)$. That means that the Hamming distance between b and c is even, contradicting the fact that $d(b,c) = 1$.

Hint for showing that if G has a Hamming labeling then G must be bipartite: If G is not bipartite, then G contains a cycle of odd length (see Section 3.4 of the Kleinberg and Tardos text). Generalize the argument given above for triangle graphs to show that this cycle cannot have a Hamming labeling. Let the length of the shortest cycle of odd length be $2k+1$ for some $k \geq 1$. Pick one node a on the cycle, and without loss of generality let $l(a) = 00..0$ (i.e., the string of all 0's). Let b and c be the two nodes of distance k from a. Consider the possibilities for the labelings of b and c.