Notes on ‘Lights Out!’

Daniel Carranza

November 26, 2020

These are my notes on the ‘Lights Out!’ puzzle for an arbitrary graph. For context, they were written before I looked up anything about this problem online.

1 The Puzzle

Suppose some light bulbs are arranged in some fashion, and these light bulbs are connected to each other arbitrarily to form a simply connected undirected graph. An example is given below.

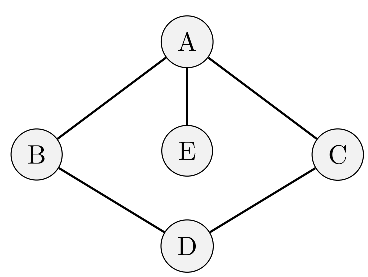

Figure 1: A possible graph of light bulbs, with each light bulb labelled.

Each of these light bulbs comes with its own switch, which may be used to turn the light bulb on or off. When flipping the switch of a light bulb, each of the light bulbs connected to it will also be affected: if a connected light bulb is off, it will turn on as well. In this example, flipping the switch corresponding to D will turn on B and C as well.

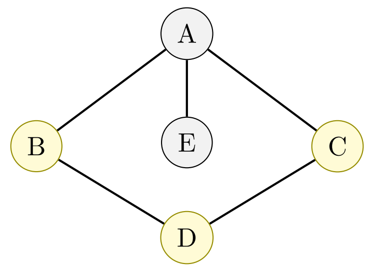

Figure 2: Flipping D's switch. This turns on B and C, which are connected to D.

However, if a connected light bulb is already on, it will turn off. In this example, after flipping D's switch, flipping A's switch will turn E on but will turn B and C off; D is unaffected.

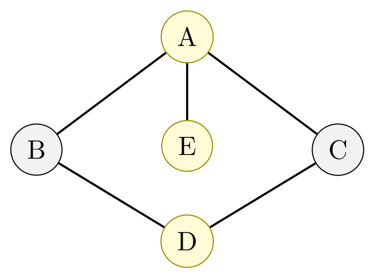

Figure 3: Flipping A's switch after flipping D's switch. Since B and C are already on after flipping D's switch, they turn off when flipping A's switch.

Naturally, we may flip the switch of a light bulb which is already on, with the same rules following. In our example, if we now flip D's switch, we get D turned off, but B and C turned back on; A is unaffected.

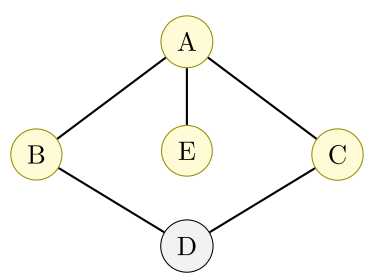

Figure 4: Flipping D's switch again. This turns off D, turns on B and C, and leaves both A and E on.

Now, given any such graph and starting with all light bulbs off, the goal is to give a sequence of flips which ends with all light bulbs turned on. In this example, a solution is to turn on D and E.

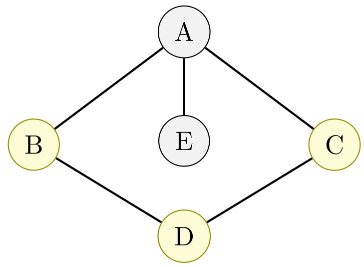

Figure 5: Flipping D on.

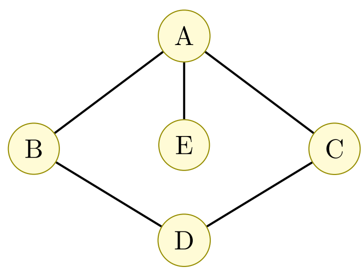

Figure 6: Flipping E on. All light bulbs are turned on, so this arrangment is solved.

2 The Analysis

We begin by fixing a graph of light bulbs. Let $L = (l_1, \ldots, l_n)$ be an ordering of the light bulbs in our graph and let $n = |L|$ be the number of light bulbs.

Definition 1. A **board state** is a configuration of each light bulb in L; either on or off. Two board states a and b are equal if for each l_i in L, l_i is off in a and off in b, or l_i is on in a and on in b.

Note that any assignment of light bulbs to on or off is a board state, even those that might not be possible to reach during the course of the puzzle. In fact, a restatement of the puzzle is to determine whether a particular board state, the board state where every light bulb is on, is possible (and, if so, determine a sequence of switches which will give this board state).

Let B be the set of all board states, including those that may or may not be possible. We may identify any board state with a vector in $(\mathbb{F}_2)^n$, where $\mathbb{F}_2$ is the field with two elements: 0 and 1. More explicitly, there is a bijection

$f\colon B \to (\mathbb{F}_2)^n$ given by

$$f(b)_i = \begin{cases} 1 & \text{if } l_i \text{ is on in } b, \\ 0 & \text{otherwise.} \end{cases}$$

Its inverse $g\colon (\mathbb{F}_2)^n \to B$ is defined as follows: given a vector $v \in (\mathbb{F}_2)^n$, the light switch l_i is on if $v_i = 1$ and off otherwise.

Proposition 1. The function $f\colon B \to (\mathbb{F}_2)^n$ as defined above is a bijection with the above inverse $g\colon (\mathbb{F}_2)^n \to B$.

Proof. We show the two composites $g \circ f = \mathrm{id}_B$ and $f \circ g = \mathrm{id}_{(\mathbb{F}_2)^n}$.

Let b be a board state and let l_i be a light bulb in L. We have that l_i is either on or off in b. If l_i is on, then $f(l_i)_i = 1$ by definition, Then, $g(f(l_i))$ turns l_i on by definition. Thus, l_i is on in b and in $(g \circ f)(b)$. If l_i is off in b, then $f(l_i) = 0$ by definition. As well, $g(f(l_i))$ turns l_i off by definition. Again, l_i is off in b and in $(g \circ f)(b)$. Therefore, $b = (g \circ f)(b)$ for any board state b, so $g \circ f = \mathrm{id}_B$.

Let v be a vector in $(\mathbb{F}_2)^n$. Given $i = 1, \ldots, n$, we have that either $v_i = 0$ or $v_i = 1$. If $v_i = 0$ then $g(v)$ turns l_i off by definition. Then, $f(g(v))_i = 0$, so that $v_i = (f \circ g)(v)_i$. Similarly, if $v_i = 1$, then $g(v)$ turns l_i on by definition. Then, $f(g(v))_i = 1$, so that $v_i = (f \circ g)(v)_i$. Then, since each component of v and $(f \circ g)(v)$ is equal, we have that $v = (f \circ g)(v)$. Therefore, $f \circ g = \mathrm{id}_{(\mathbb{F}_2)^n}$. $\square$

During the puzzle, we move from one board state to another by flipping light switches. Each light switch corresponds uniquely to a light bulb in L, so that a light switch may be identified with its respective light bulb. However, for a given light switch l_i as well as two distinct board states a and b, flipping l_i in a may result in a different board state than flipping l_i in b.

Definition 2. The **zero board state**, written 0_B, is the board state where every light bulb is off.

Proposition 2. We have that

$$f(0_B) = 0.$$

Proof. Let l_i be a light bulb in L. Since l_i is off in 0_B, we have that $f(0_B)_i = 0$. Since each component of $f(0_B)$ is zero, we get $f(0_B) = 0$. $\square$

Given a light switch l_i, we will write b_i for the board state given by flipping l_i in the zero board state. With this, we may understand the result of flipping switches on board states by analyzing their images under f. This gives us the following result:

Lemma 1. Fix a light switch l_i. Let a be a board state and b be the board state given by flipping l_i in a. We have that

$$f(b) = f(a) + f(b_i),$$

where b_i is the board state given by flipping l_i in the zero board state.

Proof. Let l_j be a light switch in L. Either l_j is connected to l_i or it is not.

Suppose l_j is connected to l_i. Then, $f(b_i) = 1$ since l_j starts out off in 0_B and is turned on by flipping l_i. Now, either l_j is on or off in a. If l_j is off in a then $f(a) = 0$. As well, since l_j is connected to l_i, it is turned on by flipping l_i so $f(b) = 1$. Then,

$$f(b) = 1 = 0 + 1 = f(a) + f(b_i).$$

Likewise, if l_j is on in a, then $f(a) = 1$. As well, since l_j is connected to l_i, it is turned off by flipping l_i so $f(b) = 0$. Then,

$$f(b) = 0 = 1 + 1 = f(a) + f(b_i).$$

In either case, we have that $f(b) = f(a) + f(b_i)$.

Now, suppose l_j is not connected to l_i. Then, $f(b_i) = 0$ since l_j starts out off in 0_B and is unaffected by flipping l_i. Either l_j is on or off in a. If l_j is off in a, then $f(a) = 0$. As well, it remains off in b, so $f(b) = 0$. Observe that

$$f(b) = 0 = 0 + 0 = f(a) + f(b_i).$$

If l_j is on in a, then $f(a) = 1$. Again, it remains on in b, so $f(b) = 1$. Now, we have that

$$f(b) = 1 = 1 + 0 = f(a) + f(b_i).$$

Again, in either case, $f(b) = f(a) + f(b_i)$.

Then, since the equality holds in the case that l_j is connected to l_i and in the case that l_j is not connected to l_i, the equality holds in general. □

This lemma gives us some useful results about flipping switches.

Corollary 1. Fix a light switch $l \in L$. Starting from any board state, flipping l twice gives the same board state as flipping it 0 times.

Proof. Fix a board state a and let b be the board state given by flipping l in 0_B. We have that

$$f(a) + f(b) + f(b) = f(a) + 2f(b) = f(a).$$

By the above lemma and the fact that f is injective, these represent the same board state as desired. □

Corollary 2. Fix light switches l_i and l_j. Starting from any board state, flipping v_i then flipping v_j gives the same board state as flipping v_j then flipping v_i.

Proof. Fix a board state a. Let b be the board state given by flipping v_i in 0_B and let c be the board state given by flipping v_j in 0_B. We have that

$$f(a) + f(b) + f(c) = f(a) + f(c) + f(b).$$

By the above lemma and the fact that f is injective, these represent the same board state as desired. □

Introducing some notation, we denote, by 1_v, the vector in $(\mathbb{F}_2)^n$ where each component is 1. Recall that each light bulb l_i defines a board state b_i by flipping its corresponding switch in the zero board state. Let $f(L)$ be the ordered set of vectors $(f(b_1), \ldots, f(b_n))$. This gives us a purely algebraic characterization of the problem.

Theorem 1. The puzzle is solvable if and only if 1_v is in the span of $f(L)$. Moreover, a linear combination $\sum_{i=0}^{n} c_i f(b_i)$ of vectors in $f(L)$ which gives $\sum_{i=0}^{n} c_i f(b_i) = 1_v$ yields a solution to the puzzle.

Proof. Suppose the puzzle is solvable. Then there is a sequence of light switches $(\lambda_i)_{i=1}^{m}$ which, starting from the zero board state, results in the board state where every light switch is on. Let us denote this board state as b. Let β_i denote the board state given by flipping λ_i in the zero board state. By Lemma 1, we have that

$$1_v = f(b) = \sum_{i=1}^{m} f(\beta_i).$$

However, by construction, $\bigcup_{i=1}^{m}\{f(\beta_i)\} \subseteq f(L)$, so we may write this as

$$1_v = f(b) = \sum_{i=1}^{n} c_i f(b_i),$$

where $c_i = 1$ if $b_i = \beta_j$ for some $j = 1, \ldots, m$ and 0 otherwise. Therefore, 1_v is in the span of $f(L)$.

Now, suppose 1_v is in the span of $f(L)$. Then, 1_v may be written as a linear combination of vectors in $f(L)$. Let us write

$$1_v = \sum_{i=0}^{n} c_i f(b_i).$$

We have that either $c_i = 0$ or $c_i = 1$. Thus, let $I \subseteq \{1, \ldots, n\}$ be the set of indices such that $c_i = 1$. This gives us that

$$1_v = \sum_{i \in I} f(b_i).$$

As well, this I induces a subset of L defined by including l_i such that $i \in I$. Then, if b denotes the board state which results from flipping this set of light switches, then Lemma 1 gives us

$$f(b) = \sum_{i \in I} f(b_i) = 1_v.$$

Then, by definition, b is the board state in which every light bulb is on. Therefore, the puzzle is solvable by flipping the switches in the I-induced subset of L. $\square$

Then, it will suffice to show that 1_v is in the span of $f(L)$. We may do this by constructing a matrix A whose columns are the elements of $f(L)$ written as column matrices. Then, 1_v is in the span of $f(B)$ if and only if $Ax = [1_v]$ is a consistent system (i.e. it has a solution). Moreover, if one can prove that this system is always consistent, then this characterization of the puzzle always gives a way to compute a solution through means such as Gaussian elimination.

To prove that this system is always consistent, we observe the following lemmas:

Lemma 2. Let A be the matrix given by writing the vectors in $f(L)$ as columns. Then, A is symmetric.

Proof. Let a_{ij} be the entry in the i-th row and j-th column of A such that $i \neq j$. By construction, this entry is 0 if flipping l_j does not affect l_i and 1 if flipping l_j also flips l_i. That is, it is 0 if l_j is connected to l_i and 1 if it is not. Now, if l_j is not connected to l_i then l_i is not connected to l_j, since our starting graph is undirected. Similarly, if l_j is connected to l_i then l_i is connected to l_j. Rephrasing this, if $a_{ij} = 0$ then $a_{ji} = 0$ and if $a_{ij} = 1$ then $a_{ji} = 1$. In either case, $a_{ij} = a_{ji}$ so A is symmetric. □

Lemma 3. Let $S \subseteq L$ be a subset of light bulbs containing an odd number of nodes. Flipping the switches corresponding to the light bulbs in S always leaves at least one light bulb in S on.

Proof. Consider the subgraph which consists of only nodes in S and edges between nodes in S. We may sum up the degrees of each node in S to get $2E$, where E is the number of edges in the subgraph; notably, $2E$ is even. Recall that an odd sum of odd numbers is always odd, so it cannot be the case that every node in S has odd degree. Then, S contains a node which is connected to an even number of nodes in S. Thus, flipping all switches in S flips l an odd number of times (it flips itself, and is flipped an even number of times for each bulb it is connected to), so it remains on after flipping all bulbs in S. □

Theorem 2. Let A be the matrix given by writing the vectors in $f(L)$ as columns, and let $[1_v]$ be the column vector where every entry is 1. Then, the system $Ax = [1_v]$ is consistent i.e. it is always solvable.

Proof. Let $A \mid [1_v]$ be the augmented matrix given by appending $[1_v]$ as a column to the right of A. Then, $Ax = [1_v]$ is consistent if and only if there is no sequence of elementary row operations on $A \mid [1_v]$ yielding a row $[0 \dots 0 \mid 1]$. Now, A is symmetric by Lemma 2, so this is identical to showing there is no sequence of elementary row operations on $A^T \mid [1_v]$ yielding the row $[0 \dots 0 \mid 1]$; notably, the rows of A^T are exactly the vectors in $f(L)$ written as columns vectors. Since swapping two rows only changes the indices and not the rows themselves and the only non-zero scalar in $\mathbb{F}_2$ is 1, it suffices to show that there is no addition of rows in $A^T \mid [1_v]$ yielding a row $[0 \dots 0 \mid 1]$.

Now, observe that since every entry in the rightmost column of $A^T \mid [1_v]$ is 1, an even sum of rows in $A^T \mid [1_v]$ always results in a 0 in the right-most entry and an odd sum of rows always results in a 1 in the right-most entry. Then, a row $[0 \dots 0 \mid 1]$ given by addition of rows in $A^T \mid [1_v]$ is exactly an odd sum of rows in A^T yielding the zero-vector; that is, an odd sum of vectors in $f(L)$ yielding the zero-vector. By Lemma 1, this corresponds to an odd number of switches in L giving the zero board state. Now, by Lemma 3, such a sequence of switches cannot exist since flipping an odd number of switches always leaves at least one light bulb on. Then, there cannot be an odd sum of rows in A^T yielding the zero-vector. □

Thus, the solution to this sytem gives us exactly a sequence of light switches to flip which results in all light bulbs being on. Thus, this result gives us a way to always construct a solution to the puzzle.

Corollary 3. The 'Lights Out!' Puzzle is solvable for any starting graph.

Proof. Fix the starting graph. By Theorem 1, this puzzle is solvable if and only if 1_v is the in the span of $f(L)$. However, Theorem 2 gives a linear combination of vectors in $f(L)$ yielding 1_v. Therefore, the puzzle is always solvable. □

3 Additional Remarks

We begin by noting that although a solution to the puzzle always exists, this solution is not always unique; there may be multiple distinct solutions to the puzzle. Under the above characterization, these solutions correspond to elements of the null-space of the associated matrix. Since an element of the null-space of A is exactly a sequence of light switches which gives the zero board state, any one solution to the puzzle, in fact, gives multiple distinct solutions; one for each sum of vectors in the null-space.

Generalizing the puzzle, we note that in general there does not exist a solution starting from an arbitrary graph state. In fact, we may give a characterization for when such a solution exists:

Proposition 3. Given a board state B, it is possible to get from B to the all-on board state if and only if it is possible to get from the zero board state to B.

Proof. Suppose it is possible to get from B to the all-on board state. This implies that we can get from the all-on board state to B, since one may simply repeat the sequence of moves which goes from B to the all-on board state. We have shown above that it is always possible to get from the zero board state to the all-on board state. By going from the zero board state to the all-on board state to B, we may get from the all-off board state to B.

Now, suppose it is possible to get from the zero board state to B. This implies that we can get from B to the zero board state, as before. Again, we know that it is possible to get from the zero board state to the all-on board state. By going from B to the zero board state to the all-on board state, it is possible to get from B to the all-on board state. □

Then, since the associated matrix is not necessarily of full rank, any board state which is not in the range of the associated matrix gives a starting board state which cannot be solved. One simple example of this is the graph K_2, the connected graph with 2 nodes, starting with one light bulb on.

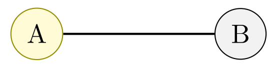

Figure 7: Here, it is impossible to turn both light bulbs on.

One interesting observation is that since a square matrix has full rank if and only if it its null-space is trivial, any graph that admits of multiple non-trivial solutions also admits initial board states which are not solvable and vice-versa.

We may also consider a variant of the puzzle which maintains the initial zero board state, but where a light bulb does not turn itself on. Again, this variant of the puzzle is not solvable for an arbitrary graph. A trivial example of this is the graph with one node.

Figure 8: Trivially, if A does not turn itself on then there is no way to turn A on.

However, we may also see that K_3 is not solvable under this ruleset.

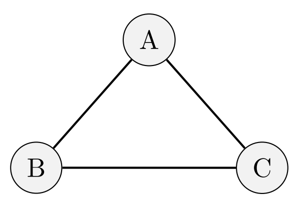

Figure 9: If each light bulb only flips its neighbours, then it is impossible to turn all light bulbs on.

Separately, allowing for directed graphs also admits of graphs which are not solvable. A concrete example of this is the following graph

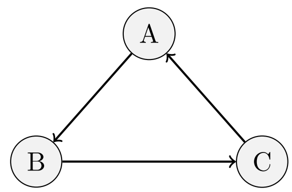

Figure 10: If a light bulb flips itself and only light bulbs in the direction of the arrows, this is a configuration where it is impossible to turn all light bulbs on.

One final generalization is one in which light bulbs may admit of additional states than “on” or “off”. Under this ruleset, flipping a light bulb may cause its state to cycle from many options, such as “off” to “red” to “blue” and back to “off”. [This is currently under investigation].