These notes are from http://math.cmu.edu/~cocox/teaching/discrete20/hats.pdf

Suppose that there are three prisoners and two hat colors. However, the guard didn't plan ahead and has only two hats of each color (so all three prisoners cannot be given the same hat color).

The exercise in the notes asked you to show that there was a deterministic strategy which guaranteed two correct guesses in this setting. Furthermore, it was mentioned that such a strategy is unique (up to relabeling the prisoners).

Firstly, here's a strategy which guarantees to save two prisoners: Prisoner $i \in [3]$ guesses the *opposite* color of prisoner $(i+1)$'s hat (where $3+1=1$). (Why does this work?)

Claim 1. *Up to relabeling the prisoners, the strategy above is unique.*

Proof. Label the two colors 0 and 1 and the prisoners $1, 2, 3$. Thus, the set of possible hat placements is

$$V = \{h \in \{0,1\}^{[3]} : h(1) + h(2) + h(3) \in \{1, 2\}\}.$$

For two hat placements h, h', define $h \triangle h' = \{i \in [3] : h(i) \neq h'(i)\}$.

Let $F = (f_1, f_2, f_3)$ be any fixed deterministic strategy. From F, we build a directed graph G_F which has vertex set V and for $h, h' \in V$, G_F has the directed edge (h, h') if and only if

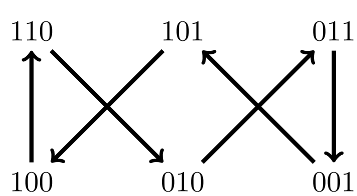
$$f_i\big((h(j))_{j \neq i}\big) = h'(i) \qquad \text{for all } i \in [3].$$

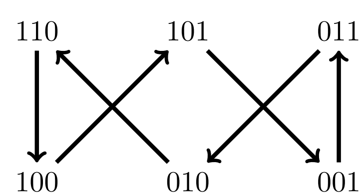
In other words, $(h, h') \in E(G_F)$ if, under the hat placement h, prisoner i guesses color $h'(i)$ for all $i \in [3]$. Observe that the strategy F is uniquely determined by the digraph G_F (why?).

Now, observe that for every $h \in V$, there is a unique $h' \in V$ for which $(h, h') \in E(G_F)$. Furthermore, if $(h, h') \in E(G_F)$, then $|h \triangle h'|$ is precisely the number of *incorrect* guesses if the guard chooses the hat placement h.

Suppose that F guarantees to save two prisoners, regardless of how the guard places hats. Thus, for any $(h, h') \in E(G_F)$, we know that $|h \triangle h'| \leq 1$. Furthermore, by the second part of the exercise in the notes, there is no way for the guard to place hats so that all three prisoners guess correctly, so in fact $|h \triangle h'| = 1$ for all $(h, h') \in E(G_F)$.

Combining these observations and the fact that $f_i\big((h(j))_{j \neq i}\big)$ does not depend on $h(i)$, we see that G_F must be one of the two following digraphs (why?):





Let $F^* = (f_1^*, f_2^*, f_3^*)$ denote the strategy mentioned above (prisoner i guesses the opposite color of prisoner $(i+1)$'s hat) and observe that G_{F^*} is precisely the digraph on the left. Thus, if $F \neq F^*$, then G_F must be the digraph on the right. However, it is then clear that $F = (f_3^*, f_2^*, f_1^*)$; that is, F is the same strategy as F^* except with the prisoners labeled differently. $\square$