You can leave your hat on (if you guess its color)

Uriel Feige
Department of Computer Science and Applied Mathematics
The Weizmann Institute
Rehovot, Israel
uriel.feige@weizmann.ac.il

March 16, 2004

Abstract

We present a methodology for solving a variety of games involving guessing the colors of hats. As an example, consider the following game. Seven players sit in a circle. There are four blue hats and four red hats. Seven hats are placed on the heads of the seven players, and the remaining hat is discarded. Every player can see the colors of the hats of the other six players, but cannot see the color of his own hat, or that of the discarded hat. Then every player needs to guess the color of his own hat. The players may coordinate a strategy before the game begins, but once the hats are placed on their heads, there is no communication of any form between the players, and in particular, no player knows whether another player already produced a guess. Is there a guessing strategy that guarantees (with absolute certainty) that at least five of the players guess correctly?

1 Introduction

In this paper we consider several mutiplayer games in which every player needs to guess the color of the hat placed on his/her head, while seeing only the colors of the hats placed on the heads of other players. In all the games that we consider, if all players guess at random, a certain number of guesses are expected to be correct. The problem is to design a deterministic strategy that guarantees that the number of correct guesses is as expected. We call such a strategy a *perfect* strategy. The trigger to the work reported here was a collection of results of Benjamin Doerr (currently unpublished), concerning variations on a setting described by Peter Winkler [?].

This paper is organized as follows. The abstract presents a puzzle that the reader may try to solve before reading the rest of the paper. In Section ?? we describe several color guessing games. In Section ?? we give some preliminary observations that may help orient the reader towards solutions to our puzzles. In Section ?? we present solutions to two of our color guessing games. In Section ?? we present a general methodology for establishing that a large class of color guessing games have perfect strategies. We show how this methodology applies to the remaining color guessing games of Section ??. In Section ?? we conclude with some additional observations and open questions.

2 The games

In all our puzzles, the number of players is denoted by n. Sometimes, n will need to be of a special form, and then we will use an additional parameter k to indicate this. For example, $n = 2k$ indicates that n is even. Players are allowed to coordinate a strategy before the game begins. The game consists of placing colored hats on the heads of the players, where C denotes the set of allowable colors in the game. For simplicity, these colors are denoted by $c_0, \dots c_{|C|-1}$. Every player can see the colors of the hat of all other players (and nothing else), and needs to guess the color of his own hat. Players have *names* (and for simplicity, the names will be $p_1, \dots, p_n$). The colors of their hats will be denoted by $h_1, \dots, h_n$. The guesses of the players are denoted by $g_1, \dots, g_n$. Formally, a strategy is a collection of n stragies, one for each player. A strategy $s_i : C^{n-1} \longrightarrow C$ for player p_i specifies the "guess" of p_i as a (deterministic) function of the tuple of colors $(h_1, \dots, h_{i-1}, h_{i+1}, \dots, h_n)$.

We may view the actual colors $h_1, \dots, h_n$ of the hats as being assigned by an *adversary* who knows the strategy of the players, and tries to select the combination of colors that causes the strategy to be least successful.

A strategy of the players will be called *perfect* if it meets some conditions (that differ from game to game). In all games, the question is to find a perfect strategy (or show that no such strategy exist).

2.1 The *plain* version

This version appears in [?].

Here $|C| = 2$ and $n = 2k$. A perfect strategy is one that guarantees that at least $k = n/2$ players correctly guess the colors of their hats.

One may also consider the case of having more than two colors for the hats. Then $|C| = c \geq 2$, $n = ck$, and at least $k = n/c$ players need to guess correctly.

2.2 The *discarded hat* version

This is the version that appears in the abstract to this paper.

Here $|C| = 2$ and $n = 4k - 1$. (The version in the abstract corresponds to the choice $k = 2$.) The players are given one more piece of information before the game begins, namely, that the total numbers of hats of color c_0 will be either $2k - 1$ or $2k$ (with the rest of the hats being of color c_1). Equivalently, one may think of there being $2k$ hats of each color, and one hat is discarded (without the players knowing which hat is discarded). In this game, a perfect strategy guarantees that at least $3k - 1$ players guess their color correctly.

2.3 The *everywhere balanced* version

Here $|C| = c \geq 2$ and n is arbitrary. The goal of the players is as follows. Let H_j be the set of the players that have a hat of color c_j. Hence $\sum_{j=0}^{c-1} |H_j| = n$. (A player does not know to which set he belongs, and the players do not know the cardinalities of the sets H_j.) A perfect strategy guarantees that in every such set H_j, the number of players who guess correctly (namely, guess c_j) is between $\lfloor |H_j|/c \rfloor$ and $\lceil |H_j|/c \rceil$.

2.4 The *majority* version

This version was studied by Doerr (private communication).

Here $|C| = 2$, n is arbitrary, and there is an additional parameter m that may depend on n. For H_j as defined above, a perfect strategy guarantees that at least $\max[|H_0|, |H_1|] - m$ players guess their color correctly. How small can m be for a perfect strategy to exist?

3 Preliminary observations

For the plain version (with $c = 2$), it is not hard to see that no strategy can guarantee more than $n/2$ correct guesses. Regardless of the strategy of the players, if the adversary assigns the colors of the hats independently at random, each player guesses his color correctly with probability half. By linearity of the expectation, the expected number of players who guess correctly is $n/2$. As expectation is an averaging operator, it follows that there is some assignment of the adversary that causes the number of correct guesses not to exceed $n/2$. (For $c > 2$, the same argument shows that no strategy can guarantee more than n/c correct guesses.)

For the discarded hat version, call the color that appears on $2k$ hats the *majority* color, and the color that appears on $2k-1$ hats the *minority* color. Every player that has a hat of the minority color sees two more hats of the majority color than the minority color, and hence knows the color of his own hat. This guarantees $2k - 1$ correct guesses. The players that have a hat of the majority color see an equal number of hats of each color. It can be shown that if the adversary assigns the colors at random, then regardless of the strategy of the players, in expectation k of the players that have a hat of the majority color guess correctly. Hence no strategy can guarantee more than $3k-1$ correct guesses.

For the everywhere balanced version, again averaging arguments show that no strategy may guarantee more than $|H_j|/c$ correct guesses, and similarly, no strategy can guarantee less than $|H_j|/c$ correct guesses. Hence the range between $\lfloor |H_j|/c \rfloor$ and $\lceil |H_j|/c \rceil$ is the best one can hope for.

For the majority version, let us bound m from below as a function of n. Recall that if the adversary assigns hat colors at random, in expectation $n/2$ players guess their color correctly. On the other hand, the majority color is expected to include $n/2+\Theta(\sqrt{n})$ hats (as this is the standard deviation of the Binomial distribution). Hence in expectation, the number of correct guesses is $\Omega(\sqrt{n})$ below $\max[|H_0|, |H_1|]$, implying that one needs $m \geq \Omega(\sqrt{n})$.

The reader may find it useful to tackle the plain version and the discarded hat version by first considering the most simple setting of the parameters, namely, that of $k = 1$. For the plain version this gives two players, one of whom needs to guess correctly. For the discarded hat version, this gives three players, two receive hats of the same color and the other receives a hat of a different color, and two players need to guess correctly.

4 Perfect strategies

In this section we present perfect strategies for two of the games of Section **??**.

4.1 The plain version

The perfect strategy presented by Winkler for this version is to pair players together (for $1 \leq i \leq k$, the pairs can be p_i and p_{i+k}), and have each pair play the case $n = 2$. That is, player p_i guesses h_{i+k} as the color of his hat, and player p_{i+k} guesses the color not equal to h_i as the color of his hat.

Another perfect strategy for the plain version (that came up in a discussion with Amir Shpilka) is based on a *symmetric* strategy. A strategy is called symmetric if the guess of a player is a function of the number of hats of each color (except his own), disregarding the distribution of hats among the players. Different players may use different symmetric functions.

For $1 \leq i \leq k$, player p_i guesses c_0 if he sees an odd number of hats of color c_0, and c_1 otherwise. For $k+1 \leq i \leq 2k = n$, player p_i guesses c_0 if he sees an even number of hats of color c_0, and c_1 otherwise. As the true total number of hats with color c_0 is either even or odd, either the first k players all guess correctly, or the last.

The above solutions generalizes easily to the case $c > 2$. We show the generalization for the symmetric strategy. Think of the color names as the numbers $0, \ldots, c-1$. For $0 \leq j \leq c-1$, for $jk+1 \leq i \leq (j+1)k$, player

i guesses for his hat the color g_i that leads to equality in $g_i + \sum_{l \neq i} h_l = j$ (modulo c). As there is some j such that $\sum_{l=1}^{n} h_l = j$ (modulo c), exactly k players guess correctly. (The above startegy in fact appears in a manuscript of Doerr for the case $c = n$.)

4.2 The discarded hat version

Recall that the difficulty in the discarded hat version is to ensure that of those players that have a hat with the majority color (that we call the *majority players*), half would guess correctly. A player call easily tell whether he is a majority player or not (but without knowing which is the majority color), and all minority players guess correctly their colors. Hence we shall only specify the strategy that players use when they are majority players.

Symmetric strategies cannot be perfect in the discarded hat version. All majority players see exactly $2k - 1$ hats of color c_0 and $2k - 1$ hats of color c_1. Hence there are only two possible symmetric strategies in this case: either guess c_0, or guess c_1. If all players have symmetric strategies, then at least $2k$ players have the same symmetric strategy (say, guess c_1). The adversary may assign these $2k$ players hats of color c_0 and all other players hats of color c_1. In this case, all majority players fail in their guess, and the number of correct guesses is only $2k - 1$.

We now present a perfect nonsymmetric strategy for the discarded hat version. Think of the players as sitting in a circle. Every player that sees $2k - 1$ hats of each color (and hence does not know the color of his own hat), computes which of the two colors occurs more times among the hats of the $(4k - 2)/2 = 2k - 1$ players that follow him in clockwise order. He then guesses the same color for his own hat.

We show now that exactly k of the majority players guess correctly. Assume without loss of generality that there are $2k$ blue hats, and consider only the players having blue hats. Starting with an arbitrary such player, name them from b_1 to b_{2k} in clockwise order. Now for every $1 \leq i \leq k$, exactly one of the two players b_i and b_{k+i} guesses correctly, because either b_{k+i} sits at most $2k - 1$ locations after b_i (and then b_i guesses correctly) or b_i sits at most $2k - 1$ locations after b_{k+i} (and then b_{k+i} guesses correctly). This is an exclusive or that follows from the fact that there are exactly $4k-1$ players altogether.

5 A general methodology

In Section **??** we presented strategies for the plain version and the discarded hat version of the color guessing game. These solutions were elegant (so the author thinks), and may well be understood by nonmathematicians. However, they were of an ad-hoc nature, and it is difficult to see how to extend them to other color guessing games (such as the everywhere balanced version and the majority version). In this section we present general principles for establishing that certain color guessing games have perfect strategies. Though we try to keep the presentation at a fairly elementary level, some level of mathematical maturity is needed in order to follow this section. In particular, we only prove the existence of a perfect strategy, but do not actually exhibit one. (Our methodology provides an algorithm for designing a perfect strategy, but the complexity of the algorithm is exponential in n. As bad as this complexity sounds, it is still much better than the time complexity of exhaustive search over all possible strategies, which is doubly exponential in n.)

A common property of all our color guessing games is that it is relatively easy to come up with randomized strategies for the players that achieve the goal of the game *in expectation*. For example, for the plain version, if every player guesses a random color for his hat, in expectation $n/2$ players guess correctly. Our methodology is based on transforming a randomized strategy into a deterministic one. Rather that discuss randomized strategies, we shall deal with what we call *fractional* strategies. In the following we shall use $\bar{h} = (h_1, \ldots, h_n) \in C^n$ to denote a vector of n hat-colors, and $\bar{h_i} = (h_1, \ldots, h_{i-1}, h_{i+1}, \ldots, h_n) \in C^{n-1}$ to denote the subvector of colors seen by player p_i.

Definition 1 *A* fractional strategy $s_i : (C, C^{n-1}) \longrightarrow C$ *for player* p_i *maps to every color* $g \in C$ *and every tuple of colors* $\bar{h_i} \in C^{n-1}$ *a value* $z_{i,g,\bar{h_i}}$, *where the following two constraints must be satisfied:*

1. *For every* i, g *and* $\bar{h}$ *we have* $z_{i,g,\bar{h_i}} \geq 0$.
2. *For every* i *and* $\bar{h}$ *we have* $\sum_{g \in C} z_{i,g,\bar{h_i}} = 1$.

Intuitively, one may view $z_{i,g,\bar{h_i}}$ as the probability that player p_i guesses color g upon seeing colors $\bar{h_i}$, or as a confidence level that p_i associates with color g upon seeing colors $\bar{h_i}$. We note that true strategies are special cases of fractional strategies, as they satisfy the additional Boolean constraint that for every $i, g, \bar{h}$, we have $z_{i,g,\bar{h_i}} \in \{0, 1\}$.

In the color guessing games, for every assignment $\bar{h}$ of hat colors, there is some goal that needs to be met by the strategies of the players (a certain number of correct guesses). We may extend this goal to apply also to fractional strategies. Recall our notation that $H_j(\bar{h})$ denotes the set of players that have a hat of color c_j in $\bar{h}$. Then for example, for the plain version, we may add the following set of constraints. For every $\bar{h}$,

$$\sum_{i \in H_0(\bar{h})} z_{i,0,\bar{h_i}} + \sum_{i \in H_1(\bar{h})} z_{i,1,\bar{h_i}} = n/2.$$

We have seen in Section **??** some perfect strategies for the plain version that indeed meet the constraint above. But assume now that we were not clever enough to find perfect strategies for the plain version. We show how our methodology can be used to infer that perfect strategies exist (but without actually exhibiting a strategy).

There is a trivial fractional strategy that satisfies all the constraints, namely, all $z_{i,g,\bar{h_i}} = 1/2$. We shall show that this strategy can be *rounded* to give a true (Boolean) strategy that still satisfies all constraints. The main point that we use in this case is that every *variable* $z_{i,g,\bar{h_i}}$ participates in exactly three constraints:

1. The *nonnegativity* constraint $z_{i,g,\bar{h_i}} \geq 0$.
2. The *strategy* constraint $\sum_{g \in C} z_{i,g,\bar{h_i}} = 1$.
3. The *goal* constraint $\sum_{j \in H_0(\bar{h})} z_{j,0,\bar{h_j}} + \sum_{j \in H_1(\bar{h})} z_{j,1,\bar{h_j}} = n/2$, for $\bar{h}$ that has $h_i = g$ and $\bar{h}$ agrees with $\bar{h_i}$ for all indices $j \neq i$.

Ignoring the nonnegativity constraints, we consider a bipartite graph (that we call the *constraint graph*) with all strategy constraints on one side, and all goal constraints on the other side. Every variable $z_{i,g,\bar{h_i}}$ contributes one edge to the bipartite graph, namely, the edge connecting the two constraints in which $z_{i,g,\bar{h_i}}$ participates. This edge is labeled by the value of $z_{i,g,\bar{h_i}}$.

We present a rounding procedure that rounds all edge values to 0/1 values, while preverving the sum of edge values incident with every vertex. This gives a true strategy for the game that satisfies the goals of the game. The rounding procedure that we describe has multiple iterations. In every iteration, some edges that previously had fractional values get integer values (either 0 or 1), and are *frozen*. When all edges are frozen, the rounding procedure ends. An invariant preserved by the rounding procedure is that for every vertex, the sum of values that are incident with it does not change.

An iteration of the rounding procedure works as follows. Assume that G contains at least one nonfrozen edge (as otherwise we are done). Observe that the subgraph induced on the nonfrozen edges has minimum degree 2, as otherwise the sum of label values incident with some vertex is noninteger. Hence it must contain a cycle. This cycle must be of even length, because the constraint graph is bipartite. Starting at an arbitrary edge e along the cycle, consider the edges of the cycle as alternating between positive and negative. Add a small value $\delta > 0$ to the values of the labels of all positive edges, and subtract δ from the values of the labels of all negative edges. Note that this does not change the sum of label values incident with any vertex. Now choose δ to be the smallest possible value that makes (at least) one such edge reach a 0/1 value. By this, at least one more edge becomes frozen and the iteration is completed.

Using the approach outlined above, one can show that both the plain version and the discarded hat version have a perfect strategy. But for the everywhere balanced game, and for the majority game, we need to address one technicality. The distinction is that in these latter games, the goal constraints are *inequalities* rather than equalities. For example, in the everywhere balanced game, the goal is to have in every set H_j between $\lfloor |H_j|/c \rfloor$ and $\lceil |H_j|/c \rceil$ correct guesses. A fractional strategy that satisfies the goal constraints might have the property that in the constraint graph, the sum of values of edges incident with a goal constraint is noninteger. If this happens, it is no longer true that in the rounding procedure the subgraph induced on nonfrozen edges has minimum degree 2. To handle this issue, we add one more vertex r on the strategy constraint side of the bipartite graph, and connect it to every vertex in the goal constraint side. Now if $s(u)$, the sum of values given by the fractional solution to the edges incident with a goal constraint u, is noninteger, then we give to the edge (r, u) the fractional value that rounds $s(u)$ up to the nearest integer. Now the sum of values incident with each goal constraint is integer. For every fractional strategy, it also must hold that the sum of values incident with each strategy constraint is integer. As the values summed up on the left hand side of the bipartite graph must be equal to the values summed up on the right hand side, it follows that also for r the sum of values incident with it is integer. Hence we can round the fractional strategy to an integer strategy. An edge (r, u) ends up with value 0 if $s(u)$ was rounded up, and with value 1 if $s(u)$ was rounded down.

Summarizing, we have the following general theorem.

Theorem 1 *Using the notation of this section, consider an arbitrary color-guessing game with the following properties:*

1. *Every goal constraint addresses one particular $\bar{h}$ and one particular set $I \subset \{1, \ldots, n\} \times C$, and has the form $q \leq \sum_{(i,g)\in I} z_{i,g,\bar{h_i}} \leq p$, where q and p are integers (possibly, $q = p$).*

2. *Every variable $z_{i,g,\bar{h_i}}$ appears in at most one goal constraint.*

Then such a game has a perfect strategy iff it has a perfect fractional strategy.

5.1 The everywhere balanced version

The goal constraints of the everywhere balanced game are for every $\bar{h}$ and $j \in \{0, \ldots, c-1\}$,

$$\lfloor |H_j|/c \rfloor \leq \sum_{i\in H_j} z_{i,c_j,\bar{h_i}} \leq \lceil |H_j|/c \rceil.$$

The conditions of Theorem **??** are satisfied. Hence we shall only show a perfect fractional strategy for the everywhere balanced game, and then Theorem **??** implies the existence of a perfect strategy (without actually exhibiting one). The fractional strategy is simple: for all $i, g, \bar{h}$ we set $z_{i,g,\bar{h_i}} = 1/c$. It is easy to check that all strategy constraints and all goal constraints are satisfied by this fractional strategy.

5.2 The majority version

Let the two colors be 0 and 1. For a given $\bar{h}$, recall that H_0 and H_1 denote the set of players with 0-hats and 1-hats respectively. For player p_i, let H_0^i

be the set of players with 0-hats that are seen by p_i (which differs from H_0 when p_i himself has a 0-hat). Consider the following fractional strategy. For a given $\bar{h}$, if $|H_0^i| \geq n/2+\sqrt{n}$, then $z_{i,0,\bar{h}_i}=1$ and $z_{i,1,\bar{h}_i}=0$. If $|H_0^i| \leq n/2-\sqrt{n}$, then $z_{i,0,\bar{h}_i}=0$ and $z_{i,1,\bar{h}_i}=1$. Else, let b denote $|H_0^i|-n/2$. Then $z_{i,0,\bar{h}_i}=1/2+b/2\sqrt{n}$ and $z_{1,1,\bar{h}_i}=1/2-b/2\sqrt{n}$.

Clearly, the above fractional strategy satisfies all strategy constraints. We now consider goal constraints. So as to slightly simplify the presentation, for every $\bar{h}$ we stregthen the corresponding goal constraint by breaking it into two separate constraints, one for H_0 and one for H_i. They are:

$$\sum_{i\in H_0} z_{i,0,\bar{h}_i} \geq \min[|H_0|, \left\lfloor |H_0|(\frac{1}{2}+\frac{|H_0|-1-n/2}{2\sqrt{n}})\right\rfloor]$$

$$\sum_{i\in H_1} z_{i,1,\bar{h}_i} \geq \min[|H_1|, \left\lfloor |H_1|(\frac{1}{2}-\frac{|H_0|-n/2}{2\sqrt{n}})\right\rfloor]$$

These constraints are satisfied by the above fractional strategy. Moreover, as goal constraints they satisfy the conditions of Theorem **??**. Hence, there is a Boolean strategy satisfying the same goal constraints.

Now, for arbitrary $\bar{h}$, if $|H_0| > n/2+\sqrt{n}$ (or $|H_1| \geq n/2+\sqrt{n}$, respectively) then all H_0 players (H_1, respectively) guess correctly and the number of correct guesses is $\max[|H_0|,|H_i|]$. If $||H_0|-|H_1|| \leq 2\sqrt{n}$, then sum up the two goal constraints. The number of correct guesses is:

$$\sum_{i\in H_0} z_{i,0,\bar{h}_i}+\sum_{i\in H_1} z_{i,1,\bar{h}_i} \geq \frac{n}{2}+(|H_0|-|H_i|)\frac{|H_0|-n/2}{2\sqrt{n}}-\frac{|H_0|}{2\sqrt{n}}-2 > \frac{n}{2}-\frac{\sqrt{n}}{2}-2.$$

As $\max[|H_0|,|H_1|] \leq n/2+\sqrt{n}$ in these cases, this gives a perfect strategy for the majority game with $m \leq 3\sqrt{n}/2$. (The leading constant in front of $\sqrt{n}$ can be improved with more careful analysis.)

6 Discussion

We presented perfect strategies for various color guessing games. The plain version and the discarded hat version can be appreciated also by nonmathematicians. The everywhere balanced and the majority version are presented so as to illustrate a general approach for solving such problems, via Theorem **??**.

In the constraint graph for the discarded hat game, all strategy constraint vertices have degree 2, and all goal constraint vertices have even degree. By shortcutting the degree 2 vertices, we get a bipartite graph of even degree, with H_0 goal constraints on one side and H_1 goal constraints on the other side. This graph is Eulerian. Every Eulerian cycle in this graph gives a perfect strategy, by having red/blue alternate along the edges of the cycle. Moreover, *every* perfect strategy for the discarded hat game can be viewed as a strategy obtained in such a way from an Eulerian cycle. (Think of the red/blue values as giving right/left orientations to the underlying edge. The resulting directed graph is Eulerian.) Using this characterization of perfect strategies it is not hard to show that every perfect strategy for the discarded hat game must be *unbiased* in the following sense: every majority player guesses *red* on exactly half of the possible hat assignments to the other players, and *blue* on the other half. (Hint: colors necessarily alternate between two successive visits to edges associated with the same player p_i.)

For the majority game, Doerr presents an explicit strategy (one for which the computation performed by every player takes time that is at most polynomial in n), but for $m = O(n^{2/3})$ rather than $m = O(\sqrt{n})$. The author does not know whether the majority game has an explicit strategy when $m = O(\sqrt{n})$.

The puzzles and solutions given in this paper are related to some well studied research areas.

Theorem **??** can be viewed as a special case of the well known fact that linear programs with integer constraints and a totally unimodular constraint matrix always have integer optimal solutions. The connection between total unimodularity and the solution of integer programs was apparently first made in [**?**], and can be found in any of a number of textbooks.

The puzzles studied here can be cast as questions about *discrepency*. The approach we used to solve them (by rounding a fractional solution) has been previously use in order to prove other results concerning discrepency, such as the well known Beck-Fiala theorem [**?**].

Acknowledgements

I thank Benjamin Doerr for communicating to me his work on games of guessing colors.

This paper is dedicated to Nadav, Ofir and Dror (ages 9, 12 and 15 at the time of the writing) who showed interest in the puzzle posed in the abstract. The color of Nadav's hat is red.

References

[1] J. Beck and T. Fiala. "Integer-making theorems". *Discrete Applied Mathematics* 3, 1–8, 1981.

[2] A. Hoffman and J. Kruskal. "Integral boundary points of convex polyhedra". In H. Kuhn and A. Tucker, editors, *Linear inequalities and related systems*, Princeton University Press, 1956.

[3] R. Newman. "You can leave your hat on". (Recorded by Joe Cocker, Tom Jones, Randy Newman.)

[4] P. Winkler. "Games people don't play". In D. Wolfe and T. Rogers, editors, *Puzzlers' tribute. A feast for the mind.* A K Peters, 2001.