T-79.4001: The Hat Problem

Janne Peltola

5.3.2008

1. The Problem
   - Introduction
   - Solutions
2. The Solution
   - Algorithm
   - Efficiency
3. Variants

The situation

n players play a game. The rules are simple.

- The players enter a room blindfolded.

The situation

n players play a game. The rules are simple.

- The players enter a room blindfolded.
- A hat is put on each player's head. The hat is either red or blue.

The situation

n players play a game. The rules are simple.

- The players enter a room blindfolded.
- A hat is put on each player's head. The hat is either red or blue.
- The players must guess the color of their hat or pass.

The situation

n players play a game. The rules are simple.

- The players enter a room blindfolded.
- A hat is put on each player's head. The hat is either red or blue.
- The players must guess the color of their hat or pass.
- The players may not communicate with each other.

The situation

n players play a game. The rules are simple.

- The players enter a room blindfolded.
- A hat is put on each player's head. The hat is either red or blue.
- The players must guess the color of their hat or pass.
- The players may not communicate with each other.
- The players win if each non-passing guess is correct.

The situation

n players play a game. The rules are simple.

- The players enter a room blindfolded.
- A hat is put on each player's head. The hat is either red or blue.
- The players must guess the color of their hat or pass.
- The players may not communicate with each other.
- The players win if each non-passing guess is correct.

We wish to find a strategy which maximizes the probability p that the players win.

Some history

- The problem was first proposed by T. Ebert (UCLA) in his Ph.D. thesis in 1998.
- There's certain charm as a puzzle but the problem doesn't seem terribly relevant.

Some history

- The problem was first proposed by T. Ebert (UCLA) in his Ph.D. thesis in 1998.
- There's certain charm as a puzzle but the problem doesn't seem terribly relevant.
- However, there are connections between its efficient algorithmic solutions for $n = 2^k - 1$ and coding theory.

Some history

- The problem was first proposed by T. Ebert (UCLA) in his Ph.D. thesis in 1998.
- There's certain charm as a puzzle but the problem doesn't seem terribly relevant.
- However, there are connections between its efficient algorithmic solutions for $n = 2^k - 1$ and coding theory.
- Thus far the problem has eluded tractable solutions for $n \neq 2^k - 1$.

3 Players: a naïve strategy

- Suppose $n = 3$. The players are Alice, Bob and Chuck.

3 Players: a naïve strategy

- Suppose $n = 3$. The players are Alice, Bob and Chuck.
- If Alice and Bob pass, Chuck guesses correctly 50% of the time.

3 Players: a naïve strategy

- Suppose $n = 3$. The players are Alice, Bob and Chuck.
- If Alice and Bob pass, Chuck guesses correctly 50% of the time.
- Can we do better than this?

3 Players: a better strategy

- Each player looks at the other players' hats.

R	R	R
R	R	B
B	B	B
B	R	R
B	B	R
R	B	B
B	R	B
R	B	R

3 Players: a better strategy

- Each player looks at the other players' hats.
- If the hats are different, he/she passes.

R	R	R
R	R	B
B	B	B
B	R	R
B	B	R
R	B	B
B	R	B
R	B	R

3 Players: a better strategy

- Each player looks at the other players' hats.
- If the hats are different, he/she passes.
- If the hats are of the same color, he/she picks the opposite one.

R	R	R
R	R	B
B	B	B
B	R	R
B	B	R
R	B	B
B	R	B
R	B	R

3 Players: a better strategy

- Each player looks at the other players' hats.
- If the hats are different, he/she passes.
- If the hats are of the same color, he/she picks the opposite one.
- This results in $p = \frac{3}{4}$, a 50% increase.

R	R	R
R	R	B
B	B	B
B	R	R
B	B	R
R	B	B
B	R	B
R	B	R

3 Players: a better strategy

- Each player looks at the other players' hats.
- If the hats are different, he/she passes.
- If the hats are of the same color, he/she picks the opposite one.
- This results in $p = \frac{3}{4}$, a 50% increase.
- This strategy can be generalized for all $n = 2^k - 1$.

R	R	R
R	R	B
B	B	B
B	R	R
B	B	R
R	B	B
B	R	B
R	B	R

7 Players

- The size of the state space is $2^7 = 128$.

7 Players

- The size of the state space is $2^7 = 128$.
- The naïve strategy can be applied.

7 Players

- The size of the state space is $2^7 = 128$.
- The naïve strategy can be applied.
- However, there is a complex strategy to achieve $p = \frac{63}{64}$.

7 Players

- The size of the state space is $2^7 = 128$.
- The naïve strategy can be applied.
- However, there is a complex strategy to achieve $p = \frac{63}{64}$.
- Next, the general strategy will be defined.

The Idea

- We have encoded certain states in C.
- We present a strategy to obtain winning sequences via C.
- We show that the strategy fails with probability $p = 2^{-r}$.

Definitions

- Consider a game of $n = 2^r - 1$ players and 2 colors.

Definitions

- Consider a game of $n = 2^r - 1$ players and 2 colors.
- Let $v = (1110100)^T \in \mathbb{Z}_2^n$ be the state vector.

Definition

Let H be an $r \times n$ matrix with entries in $\mathbb{Z}_2$. Suppose that all possible nonzero columns occur exactly once. Then H is the check matrix (ie. $Hv^T = 0$) of a code C. The code is called an $[n, n-r]$ Hamming code.

The Strategy

- Recall the state vector $v = (1110100)^T \in \mathbb{Z}_2^7$

The Strategy

- Recall the state vector $v = (1110100)^T \in \mathbb{Z}_2^7$
- We choose (for this case) $H = \begin{matrix} 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{matrix}$

The Strategy

- Recall the state vector $v = (1110100)^T \in \mathbb{Z}_2^7$
- We choose (for this case) $H = \begin{matrix} 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{matrix}$
- For person 3, there are two possible v's: $v_0 = (1100100)^T$ or $v_1 = (1110100)^T$.

The Strategy

- Recall the state vector $v = (1110100)^T \in \mathbb{Z}_2^7$
- We choose (for this case) $H = \begin{matrix} 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{matrix}$
- For person 3, there are two possible v's: $v_0 = (1100100)^T$ or $v_1 = (1110100)^T$.
- Person 3 uses the following rules to determine his call:

The Strategy

- Recall the state vector $v = (1110100)^T \in \mathbb{Z}_2^7$
- We choose (for this case) $H = \begin{matrix} 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{matrix}$
- For person 3, there are two possible v's: $v_0 = (1100100)^T$ or $v_1 = (1110100)^T$.
- Person 3 uses the following rules to determine his call:
    - If $Hv_0^T \neq 0$ and $Hv_1^T \neq 0$, he passes

The Strategy

- Recall the state vector $v = (1110100)^T \in \mathbb{Z}_2^7$
- We choose (for this case) $H = \begin{matrix} 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{matrix}$
- For person 3, there are two possible v's: $v_0 = (1100100)^T$ or $v_1 = (1110100)^T$.
- Person 3 uses the following rules to determine his call:
  - If $Hv_0^T \neq 0$ and $Hv_1^T \neq 0$, he passes
  - If $Hv_0^T = 0$ and $Hv_1^T \neq 0$, he calls Red

The Strategy

- Recall the state vector $v = (1110100)^T \in \mathbb{Z}_2^7$
- We choose (for this case) $H = \begin{matrix} 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{matrix}$
- For person 3, there are two possible v's: $v_0 = (1100100)^T$ or $v_1 = (1110100)^T$.
- Person 3 uses the following rules to determine his call:
  - If $Hv_0^T \neq 0$ and $Hv_1^T \neq 0$, he passes
  - If $Hv_0^T = 0$ and $Hv_1^T \neq 0$, he calls Red
  - If $Hv_0^T \neq 0$ and $Hv_1^T = 0$, he calls Blue

Why is this a good idea?

Lemma

There is a unique codeword c so that $d_H(v, c) = 1$.

Proof.

Existence:

- Let $s = Hv^T \neq 0$. s is a column of H, therefore $s = He_i^T$.
- $H(v + e_i)^T = s + s = 0$, so $d_H(v, c) = 1$.

$\square$

Why is this a good idea?

Lemma

There is a unique codeword c so that $d_H(v, c) = 1$.

Proof.

Uniqueness:

- Let $c_1, c_2 \in C$ so that $d_H(v, c_1) = d_H(v, c_2) = 1$.
- Therefore $v + c_1 = e_i$ and $v + c_2 = e_j$ for some (i, j).
- We have $e_i + e_j = c_1 + c_2 \in C$, so $H(e_i + e_j)^T = 0$
- $He_i^T = He_j^T$ so columns i and j must be equal, but there are no equal columns by definition $\Rightarrow$ contradiction.

$\square$

Why is this NOT a good idea?

Theorem

The strategy is successful whenever $Hv^T \neq 0$, but fails when $Hv^T = 0$.

Proof.

- We previously showed that there is a *unique* codeword c so that $d_H(v, c) = 1$.

$\square$

Why is this NOT a good idea?

Theorem

The strategy is successful whenever $Hv^T \neq 0$, but fails when $Hv^T = 0$.

Proof.

- We previously showed that there is a *unique* codeword c so that $d_H(v, c) = 1$.
- If we iterate through the vectors $v_i = v + e_i$, we find that $Hv_i^T = 0$ only for one value of i.

$\square$

Why is this NOT a good idea?

Theorem

The strategy is successful whenever $Hv^T \neq 0$, but fails when $Hv^T = 0$.

Proof.

- We previously showed that there is a *unique* codeword c so that $d_H(v, c) = 1$.
- If we iterate through the vectors $v_i = v + e_i$, we find that $Hv_i^T = 0$ only for one value of i.
- Recall that we assumed that $Hv^T \neq 0$, therefore the case $Hv_0^T = 0$ precludes all other possibilities $\Rightarrow v_0 = v$.

$\square$

Why is this NOT a good idea?

Theorem

The strategy is successful whenever $Hv^T \neq 0$, but fails when $Hv^T = 0$.

Proof.

- We previously showed that there is a *unique* codeword c so that $d_H(v, c) = 1$.
- If we iterate through the vectors $v_i = v + e_i$, we find that $Hv_i^T = 0$ only for one value of i.
- Recall that we assumed that $Hv^T \neq 0$, therefore the case $Hv_0^T = 0$ precludes all other possibilities $\Rightarrow v_0 = v$.
- If $Hv^T = 0$, no one would pass $\Rightarrow$ the strategy fails.

$\square$

A helpful lemma

Lemma

An $[n, k]$ linear code has 2^k elements.

Proof.

- An $[n, k]$ linear code C has a check matrix $H \in \mathbb{Z}_2^{k \times n}$.

$\square$

A helpful lemma

Lemma

An $[n, k]$ linear code has 2^k elements.

Proof.

- An $[n, k]$ linear code C has a check matrix $H \in \mathbb{Z}_2^{k \times n}$.
- The elements of C are the columns $+ (000)^T$

$\square$

A helpful lemma

Lemma

An $[n, k]$ linear code has 2^k elements.

Proof.

- An $[n, k]$ linear code C has a check matrix $H \in \mathbb{Z}_2^{k \times n}$.
- The elements of C are the columns $+ (000)^T$
- By combinatorics, there are 2^k elements.

$\square$

How often does one fail?

Theorem

The probability that players win by using the strategy is $1 - 2^{-r}$.

Proof.

- There are 2^n elements in $\mathbb{Z}_2^n$.

$\square$

How often does one fail?

Theorem

The probability that players win by using the strategy is $1 - 2^{-r}$.

Proof.

- There are 2^n elements in $\mathbb{Z}_2^n$.
- Therefore, the probability to find k specific elements is $\frac{2^k}{2^n}$.

□

How often does one fail?

Theorem

The probability that players win by using the strategy is $1 - 2^{-r}$*.*

Proof.

- There are 2^n elements in $\mathbb{Z}_2^n$.
- Therefore, the probability to find k specific elements is $\frac{2^k}{2^n}$.
- Therefore, the probability to hit an element v so that $Hv^T = 0$ is 2^{-r}.

$\square$

How often does one fail?

Theorem

The probability that players win by using the strategy is $1 - 2^{-r}$.

Proof.

- There are 2^n elements in $\mathbb{Z}_2^n$.
- Therefore, the probability to find k specific elements is $\frac{2^k}{2^n}$.
- Therefore, the probability to hit an element v so that $Hv^T = 0$ is 2^{-r}.
- The complement case's $(Hv^T \neq 0)$ probability is $1 - 2^{-r}$.

$\square$

Food for thought

- In the best tradition of the literature, I explore the opportunities for expanding the problem without providing actual solutions.
- One could consider k colors instead of 2.

Food for thought

- In the best tradition of the literature, I explore the opportunities for expanding the problem without providing actual solutions.
- One could consider k colors instead of 2.
- One could consider $n \neq 2^r - 1$.

Food for thought

- In the best tradition of the literature, I explore the opportunities for expanding the problem without providing actual solutions.
- One could consider k colors instead of 2.
- One could consider $n \neq 2^r - 1$.
- One could consider non-symmetrical color distributions.

Food for thought

- In the best tradition of the literature, I explore the opportunities for expanding the problem without providing actual solutions.
- One could consider k colors instead of 2.
- One could consider $n \neq 2^r - 1$.
- One could consider non-symmetrical color distributions.
- One could consider variants with more error tolerance (i.e. 1 error allowed). Could one solve $n \neq 2^r - 1$ with relaxed conditions?

Summary

- We found an interesting connection between coding theory and game theory.
- One can encode certain states and then make sure that there are only unique legal states which encode to nearby codewords.
- One can use this knowledge to develop a near-optimal strategy.