The Penney's Game with Group Action

2020 PRIMES Conference

Sean Li
Mentor: Tanya Khovanova

October 17-18, 2020

Coin flipping problem

Example (classic)

I repeatedly flip a coin. What is the expected number of flips for me to flip two heads in a row?

Coin flipping problem

Example (classic)

I repeatedly flip a coin. What is the expected number of flips for me to flip two heads in a row?

- Markov chains with states $\varnothing, H$ and end state HH to get EV of $\boxed{6}$.

Coin flipping problem

Example (classic)

I repeatedly flip a coin. What is the expected number of flips for me to flip two heads in a row?

- Markov chains with states $\varnothing$, H and end state HH to get EV of $\boxed{6}$.

Modification

What about HT instead of HH?

Coin flipping problem

Example (classic)

I repeatedly flip a coin. What is the expected number of flips for me to flip two heads in a row?

- Markov chains with states $\varnothing, H$ and end state HH to get EV of $\boxed{6}$.

Modification

What about HT instead of HH?

- Answer of $\boxed{4}$; intuitively less since HT has no "reset."

Coin flipping problem

Example (classic)

I repeatedly flip a coin. What is the expected number of flips for me to flip two heads in a row?

- Markov chains with states $\varnothing, H$ and end state HH to get EV of $\boxed{6}$.

Modification

What about HT instead of HH?

- Answer of $\boxed{4}$; intuitively less since HT has no "reset."

General question

A random string is generated by attaching either $\mathtt{H}$ or $\mathtt{T}$ to the end of the string until the substring ___ appears. What is the expected length of the final string?

- In other words, calculate the **expected wait time**.

Natural questions

- Why does HH take longer than HT? Can we generalize?

Natural questions

- Why does HH take longer than HT? Can we generalize?
- More letters in the alphabet?

Problem (Rubinstein-Salzedo)

A monkey hits one of 26 letters at random each second. On average, how long will it take for the monkey to type the word ABRACADABRA?

Natural questions

- Why does HH take longer than HT? Can we generalize?
- More letters in the alphabet?

Problem (Rubinstein-Salzedo)

A monkey hits one of 26 letters at random each second. On average, how long will it take for the monkey to type the word ABRACADABRA?

- How many outputs of length ℓ do not contain a word except at the end?

Natural questions

- Why does HH take longer than HT? Can we generalize?
- More letters in the alphabet?

Problem (Rubinstein-Salzedo)

A monkey hits one of 26 letters at random each second. On average, how long will it take for the monkey to type the word ABRACADABRA?

- How many outputs of length ℓ do not contain a word except at the end?
- What if we stop when we see any word from a set S? (e.g. $S = \{HHHH, THHH\}$, leading to the concept of the **Penney's game**)

Basic terminology

We designate an **alphabet** $\mathcal{A}$ with q letters.

Example

When flipping a coin, $q = 2$ and $\mathcal{A} = \{\texttt{H}, \texttt{T}\}$, the *coin alphabet*.

Basic terminology

We designate an **alphabet** $\mathcal{A}$ with q letters.

Example

When flipping a coin, $q = 2$ and $\mathcal{A} = \{\mathtt{H}, \mathtt{T}\}$, the *coin alphabet*.

A **word** w *avoids* v if it does not contain any substring equal to v. We are interested in:

- words which avoid w, except for a single w at the end; and
- words which avoid a word w.

Word correlation

The **autocorrelation** of a word $w = w_1 w_2 \dots w_\ell$ of length ℓ is a vector

$$C(w, w) = (C_0, C_1, \dots, C_{\ell-1})$$

of 0's and 1's, such that $C_k = 1$ iff w has period k, i.e. $w_i = w_{i+k}$.

Word correlation

The **autocorrelation** of a word $w = w_1 w_2 \ldots w_\ell$ of length ℓ is a vector

$$C(w, w) = (C_0, C_1, \ldots, C_{\ell-1})$$

of 0's and 1's, such that $C_k = 1$ iff w has period k, i.e. $w_i = w_{i+k}$.

Example

Under the coin alphabet, the word $w = \texttt{HTHTH}$ has autocorrelation $C(w, w) = (1, 0, 1, 0, 1)$.

Word correlation

The **autocorrelation** of a word $w = w_1 w_2 \ldots w_\ell$ of length ℓ is a vector

$$C(w,w) = (C_0, C_1, \ldots, C_{\ell-1})$$

of 0's and 1's, such that $C_k = 1$ iff w has period k, i.e. $w_i = w_{i+k}$.

Example

Under the coin alphabet, the word $w = \texttt{HTHTH}$ has autocorrelation $C(w,w) = (1,0,1,0,1)$.

$$\begin{array}{l}
\underline{\texttt{HTHTH}} \\
\texttt{HTHTH} \qquad\quad \to 1 \\
\;\;\texttt{HTHT H} \qquad \to 0 \\
\;\;\;\;\texttt{HTH TH} \quad\; \to 1 \\
\;\;\;\;\;\;\texttt{HT HTH} \;\; \to 0 \\
\;\;\;\;\;\;\;\;\texttt{H THTH} \to 1
\end{array}$$

Word correlation (pt. 2)

Similarly, we can define the **correlation** $C(w, v)$ between two words w and v, even between words of different length.

Word correlation (pt. 2)

Similarly, we can define the **correlation** $C(w, v)$ between two words w and v, even between words of different length.

Example

Under the coin alphabet, the correlation $C(\texttt{HTHTTH}, \texttt{HTTHT})$ is equal to $(0, 0, 1, 0, 0, 1)$.

$$
\begin{array}{ll}
\underline{\texttt{HTHTTH}} & \\
\texttt{HTTHT} & \to 0 \\
\;\texttt{HTTHT} & \to 0 \\
\;\;\texttt{HTTH T} & \to 1 \\
\;\;\;\texttt{HTT HT} & \to 0 \\
\;\;\;\;\texttt{HT THT} & \to 0 \\
\;\;\;\;\;\texttt{H TTHT} & \to 1
\end{array}
$$

Expected wait times

Say $C(w,w) = (C_0, C_1, \dots, C_{\ell-1})$. The **Conway leading number** $w\mathrm{L}w$ of a word w of length ℓ is

$$C_0 q^{\ell-1} + C_1 q^{\ell-2} + \cdots + C_{\ell-1}.$$

(Think base-q.)

Expected wait times

Say $C(w,w) = (C_0, C_1, \ldots, C_{\ell-1})$. The **Conway leading number** $w\mathrm{L}w$ of a word w of length ℓ is

$$C_0 q^{\ell-1} + C_1 q^{\ell-2} + \cdots + C_{\ell-1}.$$

(Think base-q.)

Theorem (Collings, 1982)

The expected wait time for a word w is precisely $q \cdot w\mathrm{L}w$.

Expected wait times

Say $C(w,w) = (C_0, C_1, \ldots, C_{\ell-1})$. The **Conway leading number** $w\mathrm{L}w$ of a word w of length ℓ is

$$C_0 q^{\ell-1} + C_1 q^{\ell-2} + \cdots + C_{\ell-1}.$$

(Think base-q.)

Theorem (Collings, 1982)

The expected wait time for a word w is precisely $q \cdot w\mathrm{L}w$.

Example

On average, it takes $2(2^{4-0} + 2^{4-2} + 2^{4-4}) = 42$ letters to generate the string HTHTH.

Expected wait times

Say $C(w,w) = (C_0, C_1, \ldots, C_{\ell-1})$. The **Conway leading number** $w\mathrm{L}w$ of a word w of length ℓ is

$$C_0 q^{\ell-1} + C_1 q^{\ell-2} + \cdots + C_{\ell-1}.$$

(Think base-q.)

Theorem (Collings, 1982)

The expected wait time for a word w is precisely $q \cdot w\mathrm{L}w$.

Example

On average, it takes $2(2^{4-0} + 2^{4-2} + 2^{4-4}) = 42$ letters to generate the string HTHTH.

Example (Rubinstein-Salzedo)

On average, a monkey will take $26(26^{10} + 26^3 + 1) = 26^{11} + 26^4 + 26$ letters to type ABRACADABRA.

The Penney's game

If the avoiding set S has two words $\{w_A, w_B\}$, then we can turn this into a game (the **Penney's game**, a la Walter Penney): if Alice and Bob pick w_A and w_B, which word will appear first?

The Penney's game

If the avoiding set S has two words $\{w_A, w_B\}$, then we can turn this into a game (the **Penney's game**, a la Walter Penney): if Alice and Bob pick w_A and w_B, which word will appear first?

The game is **non-transitive**: take the cycle

$$\mathtt{HHT} < \mathtt{THH} < \mathtt{TTH} < \mathtt{HTT} < \mathtt{HHT}.$$

The Penney's game

If the avoiding set S has two words $\{w_A, w_B\}$, then we can turn this into a game (the **Penney's game**, a la Walter Penney): if Alice and Bob pick w_A and w_B, which word will appear first?

The game is **non-transitive**: take the cycle

$$\mathtt{HHT} < \mathtt{THH} < \mathtt{TTH} < \mathtt{HTT} < \mathtt{HHT}.$$

Theorem (Conway)

The odds that Alice wins are exactly

$$p_A : p_B = (w_B \mathsf{L} w_B - w_B \mathsf{L} w_A) : (w_A \mathsf{L} w_A - w_A \mathsf{L} w_B).$$

The Penney's game

If the avoiding set S has two words $\{w_A, w_B\}$, then we can turn this into a game (the **Penney's game**, a la Walter Penney): if Alice and Bob pick w_A and w_B, which word will appear first?

The game is **non-transitive**: take the cycle

$$\mathtt{HHT} < \mathtt{THH} < \mathtt{TTH} < \mathtt{HTT} < \mathtt{HHT}.$$

Theorem (Conway)

The odds that Alice wins are exactly

$$p_A : p_B = (w_B \mathsf{L} w_B - w_B \mathsf{L} w_A) : (w_A \mathsf{L} w_A - w_A \mathsf{L} w_B).$$

Example

If Alice picks HTHT and Bob picks THTT, then Alice's chance of winning is $\frac{9}{14}$. Alice's expected wait time is 20, but Bob's is 18.

Best beater

Theorem (Guibas & Odlyzko)

If Alice picks her word $w_A = w(1)w(2)\ldots w(\ell)$ first, then Bob has the best odds of winning when he chooses $w_B = w^* w(1)w(2)\ldots w(\ell-1)$ for some w^*. This is a winning strategy.

Best beater

Theorem (Guibas & Odlyzko)

If Alice picks her word $w_A = w(1)w(2)\dots w(\ell)$ *first, then Bob has the best odds of winning when he chooses* $w_B = w^* w(1)w(2)\dots w(\ell-1)$ *for some* w^*. *This is a winning strategy.*

Example

If Alice chooses HHHH and Bob chooses THHH, then the probability Bob's word appears first is $\frac{15}{16}$.

Best beater (pt. 2)

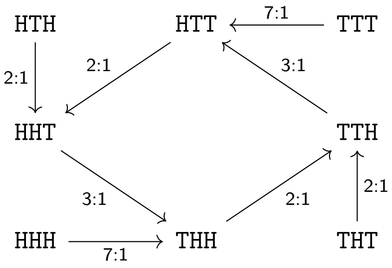

Figure: Directed graph of best beaters for $(q, \ell) = (2, 3)$.

Generating function system

For a word w, define

- the **correlation polynomial**

$$C_{v,w}(z) = C_0 + C_1 z + \cdots + C_{\ell-1} z^{\ell-1}.$$

Generating function system

For a word w, define

- the **correlation polynomial**

$$C_{v,w}(z) = C_0 + C_1 z + \cdots + C_{\ell-1} z^{\ell-1}.$$

- $T_w(n)$ be the number of words of length n avoiding w, except for a single w at the end; and
- $A_w(n)$ be the number of words of length n avoiding w.

Generating function system

For a word w, define

- the **correlation polynomial**

$$C_{v,w}(z) = C_0 + C_1 z + \cdots + C_{\ell-1} z^{\ell-1}.$$

- $T_w(n)$ be the number of words of length n avoiding w, except for a single w at the end; and
- $A_w(n)$ be the number of words of length n avoiding w.

We then define the generating functions

$$G(z) = \sum_{n=0}^{\infty} A_w(n) z^n, \quad G_w(z) = \sum_{n=0}^{\infty} T_w(n) z^n.$$

Extended results

Theorem (Guibas & Odlyzko, 1978)

For a **reduced** *set* $S = \{w_1, w_2, \ldots, w_k\}$, *the generating functions* $G(z)$, $G_{w_1}(z)$, $G_{w_2}(z)$, $\ldots$, $G_{w_k}(z)$ *satisfy the following system of linear equations:*

$$(1 - qz)G(z) + G_{w_1}(z) + G_{w_2}(z) + \cdots + G_{w_k}(z) = 1$$

$$G(z) - z^{-\ell_1} C_{w_1,w_1}(z) G_{w_1}(z) - \cdots - z^{-\ell_k} C_{w_k,w_1}(z) G_{w_k}(z) = 0$$

$$G(z) - z^{-\ell_1} C_{w_1,w_2}(z) G_{w_1}(z) - \cdots - z^{-\ell_k} C_{w_k,w_2}(z) G_{w_k}(z) = 0$$

$$\vdots$$

$$G(z) - z^{-\ell_1} C_{w_1,w_k}(z) G_{w_1}(z) - \cdots - z^{-\ell_k} C_{w_k,w_k}(z) G_{w_k}(z) = 0$$

Patterns

Motivating idea is generalizing the game: what if Alice and Bob pick sets of words? What if Alice can pick abaa, representing `HTHH` and `THTT` at the same time?

Patterns

Motivating idea is generalizing the game: what if Alice and Bob pick sets of words? What if Alice can pick abaa, representing `HTHH` and `THTT` at the same time?

We generalize to a group action $\varphi : G \times \mathcal{A} \to \mathcal{A}$, thus sending words to words. A word w resides in an orbit $\mathcal{O}(w)$ under the action; a **pattern** is a representative (typically earliest lex. and *in lowercase*) from $\mathcal{O}(w)$. Alice and Bob pick patterns instead of words.

Patterns

Motivating idea is generalizing the game: what if Alice and Bob pick sets of words? What if Alice can pick abaa, representing HTHH and THTT at the same time?

We generalize to a group action $\varphi : G \times \mathcal{A} \to \mathcal{A}$, thus sending words to words. A word w resides in an orbit $\mathcal{O}(w)$ under the action; a **pattern** is a representative (typically earliest lex. and *in lowercase*) from $\mathcal{O}(w)$. Alice and Bob pick patterns instead of words.

Example

Under the alphabet $\mathcal{A} = \{\mathtt{A}, \mathtt{B}, \mathtt{C}\}$ and the cyclic action, the orbit of ABC is $\{\mathtt{ABC}, \mathtt{BCA}, \mathtt{CAB}\}$ and corresponds to the pattern abc.

Patterns

Motivating idea is generalizing the game: what if Alice and Bob pick sets of words? What if Alice can pick abaa, representing HTHH and THTT at the same time?

We generalize to a group action $\varphi : G \times \mathcal{A} \to \mathcal{A}$, thus sending words to words. A word w resides in an orbit $\mathcal{O}(w)$ under the action; a **pattern** is a representative (typically earliest lex. and *in lowercase*) from $\mathcal{O}(w)$. Alice and Bob pick patterns instead of words.

Example

Under the alphabet $\mathcal{A} = \{\mathtt{A}, \mathtt{B}, \mathtt{C}\}$ and the cyclic action, the orbit of ABC is $\{\mathtt{ABC}, \mathtt{BCA}, \mathtt{CAB}\}$ and corresponds to the pattern abc.

We can also define pattern correlation polynomials, and generating functions $G(z)$ with avoiders of p, and $G_p(z)$ for first-timers of p.

How the Penney flips

- If w is in the orbit represented by p, then

$$\mathcal{C}(p,p) = \sum_{v \in \mathcal{O}(w)} C(v,w).$$

How the Penney flips

- If w is in the orbit represented by p, then

$$\mathcal{C}(p,p) = \sum_{v \in \mathcal{O}(w)} C(v,w).$$

- Generally, if w' is in the orbit represented by p', and w in the orbit of p, then

$$\mathcal{C}(p,p') = \sum_{v \in \mathcal{O}(w)} C(v,w').$$

How the Penney flips

- If w is in the orbit represented by p, then

$$\mathcal{C}(p,p) = \sum_{v \in \mathcal{O}(w)} C(v,w).$$

- Generally, if w' is in the orbit represented by p', and w in the orbit of p, then

$$\mathcal{C}(p,p') = \sum_{v \in \mathcal{O}(w)} C(v,w').$$

- The pattern Conway leading number and correlation polynomial

$$p\mathcal{L}p' = \mathcal{C}_0 q^{\ell-1} + \mathcal{C}_1 q^{\ell-2} + \cdots + \mathcal{C}_{\ell-1};$$
$$\mathcal{C}_{p,p'}(z) = \mathcal{C}_0 + \mathcal{C}_1 z + \cdots + \mathcal{C}_{\ell-1} z^{\ell-1}.$$

How the Penney flips (pt. 2)

Theorem

For a **reduced** *set* $S = \{p_1, p_2, \ldots, p_k\}$ *whose orbit sizes are* $r_1, r_2, \ldots, r_k$*, the generating functions* $G(z)$*,* $G_{p_1}(z)$*,* $G_{p_2}(z)$*, ...,* $G_{p_k}(z)$ *satisfy the following system of linear equations:*

$$(1 - qz)G(z) + G_{p_1}(z) + G_{p_2}(z) + \cdots + G_{p_k}(z) = 1$$

$$G(z) - \frac{1}{r_1} z^{-\ell_1} \mathcal{C}_{p_1,p_1}(z) G_{p_1}(z) - \cdots - \frac{1}{r_k} z^{-\ell_k} \mathcal{C}_{p_k,p_1}(z) G_{p_k}(z) = 0$$

$$G(z) - \frac{1}{r_1} z^{-\ell_1} \mathcal{C}_{p_1,p_2}(z) G_{p_1}(z) - \cdots - \frac{1}{r_k} z^{-\ell_k} \mathcal{C}_{p_k,p_2}(z) G_{p_k}(z) = 0$$

$$\vdots$$

$$G(z) - \frac{1}{r_1} z^{-\ell_1} \mathcal{C}_{p_1,p_k}(z) G_{p_1}(z) - \cdots - \frac{1}{r_k} z^{-\ell_k} \mathcal{C}_{p_k,p_k}(z) G_{p_k}(z) = 0$$

How the Penney flips (pt. 3)

- All results from words carry over to patterns (expected value, odds, winning strategy for sufficiently long words)

How the Penney flips (pt. 3)

- All results from words carry over to patterns (expected value, odds, winning strategy for sufficiently long words)

Theorem

Suppose Alice picks the pattern p_A and Bob picks the pattern p_B. The odds that Alice wins are exactly

$$\frac{1}{r_B}(p_B \mathcal{L} p_B - p_B \mathcal{L} p_A) : \frac{1}{r_A}(p_A \mathcal{L} p_A - p_A \mathcal{L} p_B).$$

(Here r_A and r_B are orbit sizes.)

How the Penney flips (pt. 3)

- All results from words carry over to patterns (expected value, odds, winning strategy for sufficiently long words)

Theorem

Suppose Alice picks the pattern p_A and Bob picks the pattern p_B. The odds that Alice wins are exactly

$$\frac{1}{r_B}(p_B \mathcal{L} p_B - p_B \mathcal{L} p_A) : \frac{1}{r_A}(p_A \mathcal{L} p_A - p_A \mathcal{L} p_B).$$

(Here r_A and r_B are orbit sizes.)

- Cyclic action (using $G = C_q$) and symmetric action (using $G = S_q$)

How the Penney flips (pt. 3)

- All results from words carry over to patterns (expected value, odds, winning strategy for sufficiently long words)

Theorem

Suppose Alice picks the pattern p_A and Bob picks the pattern p_B. The odds that Alice wins are exactly

$$\frac{1}{r_B}(p_B \mathcal{L} p_B - p_B \mathcal{L} p_A) : \frac{1}{r_A}(p_A \mathcal{L} p_A - p_A \mathcal{L} p_B).$$

(Here r_A and r_B are orbit sizes.)

- Cyclic action (using $G = C_q$) and symmetric action (using $G = S_q$)

Theorem

For lengths $\ell < q - \sqrt{q}$ and under the symmetric action, Alice has a winning strategy by choosing a pattern with ℓ distinct letters $a_1 a_2 \ldots a_\ell$.

How the Penney flips (pt. 4)

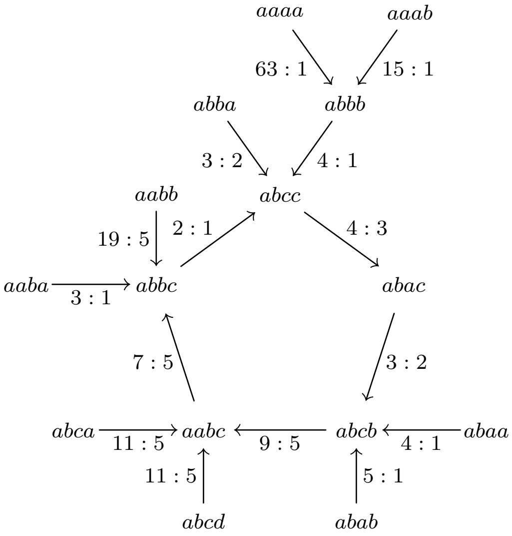

Figure: Directed graph of Bob's best choices for $(q, \ell) = (4, 4)$.

Acknowledgements

- My mentor, Dr. Tanya Khovanova of MIT
- The PRIMES Program, especially Dr. Gerovitch & Dr. Etingof
- Peers at MIT-PRIMES and the attendees of MathROCs
- My parents :)

References

- Elwyn R. Berlekamp, John H. Conway and Richard K. Guy, *Winning Ways for your Mathematical Plays*, 2nd Edition, Volume 4, AK Peters, 2004, 885.
- Stanley Collings. Coin Sequence Probabilities and Paradoxes, *Inst. Math A*, **18** (1982), 227–232.
- L.J. Guibas and A.M. Odlyzko. String Overlaps, Pattern Matching, and Nontransitive Games, *J. Comb. Theory A*, **30** (1981), no. 2, 183–208.
- Walter Penney, Problem 95. Penney-Ante, *J. Recreat. Math.*, **2** (1969), 241.