Penney's game paradoxes

Wanda Niemyska

Institute of Mathematics
University of Silesia
Poland

May 11th, 2012

ISCAMI, Czech Republic, May 2012

References

[1] W. Penney: *Problem: penney-ante*, Journal of Recreational Mathematics, 2 (1969), 241.

[2] W. Niemyska: *O pewnych zagadnieniach związanych z grami Penneya* (in Polish), Warsaw Uniwersity, Masters Thesis (2010)

[3] J. Jakubowski, R. Sztencel: *Wstęp do teorii prawdopodobieństwa* (in Polish), (2000), 273-274, 438-439.

[4] Shuo-Yen Robert Li: *A Martingale Approach to the Study of Occurence of Sequence Patterns In Repeated Experiments*, Annals of Probability, 8 (1980), 1171-1176

[5] Leonidas J. Guibas, Andrew M. Odlyzko: *String Overlaps, Pattern Matching, and Nontransitive Games*, Journal of Combinatorial Theory, ser. A, vol. 30 (1981), 183-208

Rules of the game

Rules of the classic version of Penney's game:

- 2 players;
- each player selects a three-bit long sequence of heads and tails;
- a coin is tossed until one of those sequences appears as a subsequence of the coin toss outcomes;
- the coin is fair;
- the player whose sequence appears first wins.

Rules of the game

Rules of the classic version of Penney's game:

- 2 players;
- each player selects a three-bit long sequence of heads and tails;
- a coin is tossed until one of those sequences appears as a subsequence of the coin toss outcomes;
- the coin is fair;
- the player whose sequence appears first wins.

There are 8 different three-bit long sequences of heads and tails:
HHH, HHT, HTH, THH, HTT, THT, TTH, TTT

Rules of the game

Rules of the classic version of Penney's game:

- 2 players;
- each player selects a three-bit long sequence of heads and tails;
- a coin is tossed until one of those sequences appears as a subsequence of the coin toss outcomes;
- the coin is fair;
- the player whose sequence appears first wins.

There are 8 different three-bit long sequences of heads and tails:
HHH, HHT, HTH, THH, HTT, THT, TTH, TTT

$\binom{8}{2} = 28$ classic games

Example of the game

Example of the game

Do the players have the same chances of winning in this game?

Stochastic graph's definition

We call a *stochastic graph* an ordered pair (S, Q), where:

1. S - a finite set of vertices;
2. Q - a function from SxS into $\mathbb{R}$ such that the following conditions are satisfied:
   - $Q(i,j) \geq 0 \quad$ for all $i, j \in S$;
   - $\sum_{j \in S} Q(i,j) = 1 \quad$ for all $i \in S$.

Stochastic graph's definition

We call a *stochastic graph* an ordered pair (S, Q), where:

1. S - a finite set of vertices;
2. Q - a function from SxS into $\mathbb{R}$ such that the following conditions are satisfied:
   - $Q(i,j) \geq 0 \quad$ for all $i, j \in S$;
   - $\sum_{j \in S} Q(i,j) = 1 \quad$ for all $i \in S$.

- $Q(i,j)$ is the probability of transition between states i and j in one step. We use the notation p_{ij}.

Stochastic graph's definition

We call a *stochastic graph* an ordered pair (S, Q), where:

1. S - a finite set of vertices;
2. Q - a function from SxS into $\mathbb{R}$ such that the following conditions are satisfied:
   - $Q(i,j) \geq 0 \quad$ for all $i, j \in S$;
   - $\sum_{j \in S} Q(i,j) = 1 \quad$ for all $i \in S$.

- $Q(i,j)$ is the probability of transition between states i and j in one step. We use the notation p_{ij}.
- A pair (i,j) is an edge of the graph, if $p_{ij} > 0$.

Stochastic graph's definition

We call a *stochastic graph* an ordered pair (S, Q), where:

1. S - a finite set of vertices;
2. Q - a function from SxS into $\mathbb{R}$ such that the following conditions are satisfied:
   - $Q(i,j) \geq 0 \quad$ for all $i, j \in S$;
   - $\sum_{j \in S} Q(i,j) = 1 \quad$ for all $i \in S$.

- $Q(i,j)$ is the probability of transition between states i and j in one step. We use the notation p_{ij}.
- A pair (i,j) is an edge of the graph, if $p_{ij} > 0$.
- If $p_{jj} = 1$, then j is a boundary vertice.

Stochastic graph's definition

We call a *stochastic graph* an ordered pair (S, Q), where:

1. S - a finite set of vertices;
2. Q - a function from SxS into $\mathbb{R}$ such that the following conditions are satisfied:
   - $Q(i,j) \geq 0 \quad$ for all $i, j \in S$;
   - $\sum_{j \in S} Q(i,j) = 1 \quad$ for all $i \in S$.

- $Q(i,j)$ is the probability of transition between states i and j in one step. We use the notation p_{ij}.
- A pair (i,j) is an edge of the graph, if $p_{ij} > 0$.
- If $p_{jj} = 1$, then j is a boundary vertice.
- The set B of all boundary vertices is called the boundary of the graph.

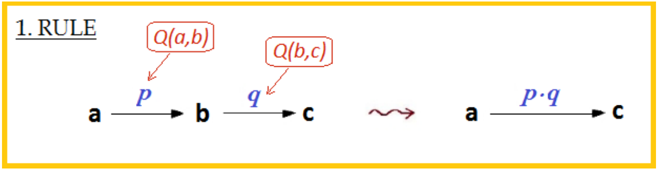
The rules of reduction

1. RULE

Q(a,b) Q(b,c)

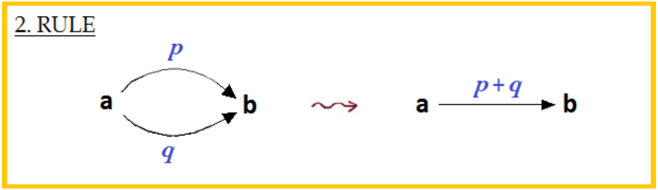
$$a \xrightarrow{p} b \xrightarrow{q} c \quad \rightsquigarrow \quad a \xrightarrow{p\cdot q} c$$

The rules of reduction



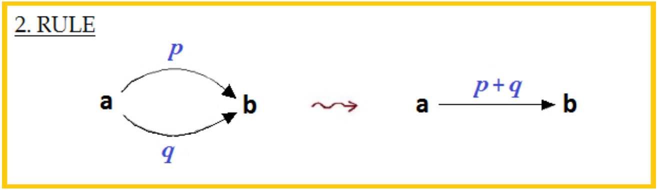
$$P(A \cap B) = P(A) \cdot P(B), \quad A, B - \text{independent events}$$

The rules of reduction



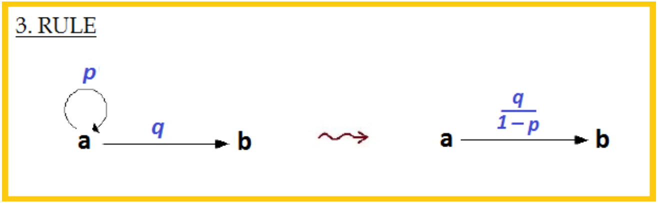
$$P(A \cup B) = P(A) + P(B), \quad A, B - \text{mutually exclusive events}$$

The rules of reduction



$$P(A \cup B) = P(A) + P(B), \quad A, B - \text{mutually exclusive events}$$

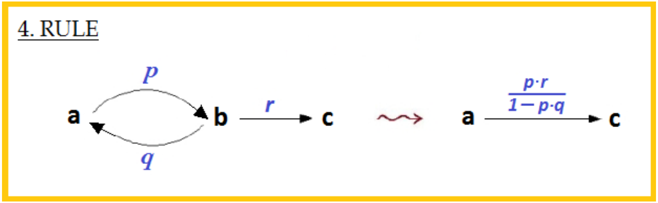
The rules of reduction

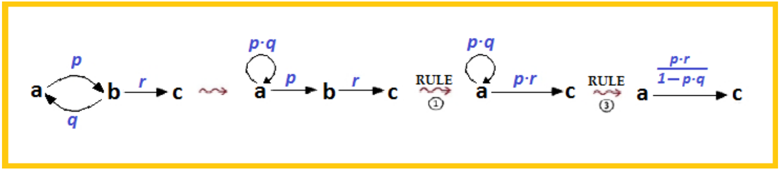


$$a \to b, \qquad a \to a \to b, \qquad a \to a \to a \to b, \ldots \qquad \Rightarrow$$

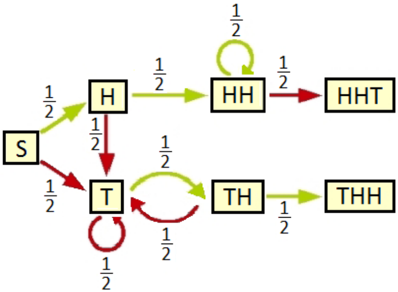
$$q + p \cdot q + p \cdot p \cdot q + \ldots = q \cdot \frac{1}{1-p}$$

The rules of reduction

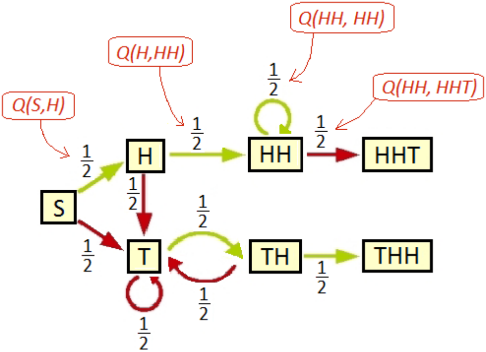




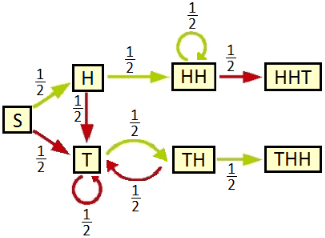
Game HHT-THH



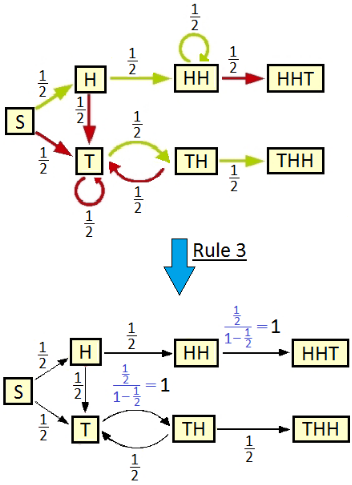
Game HHT-THH



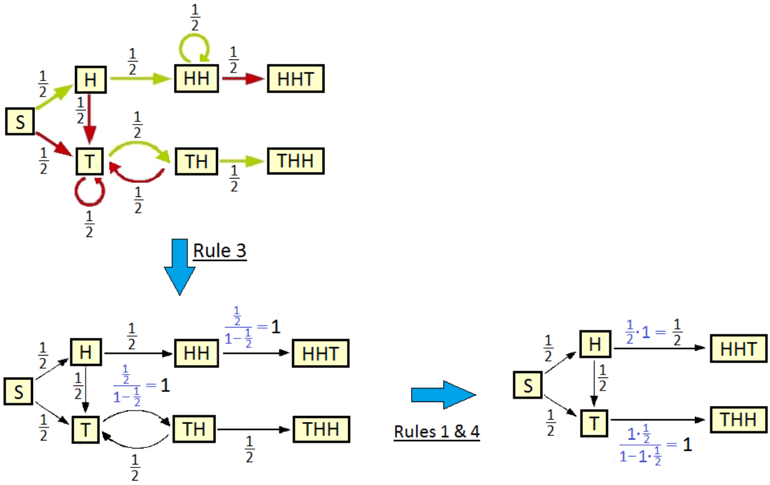
Game HHT-THH



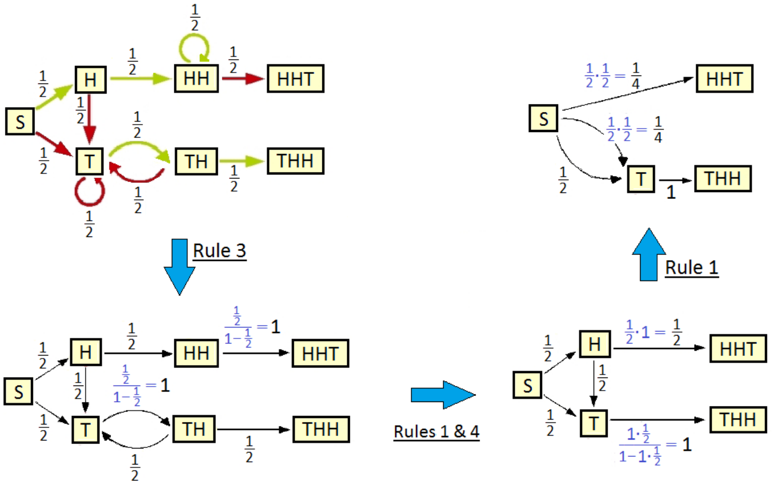
Game HHT-THH



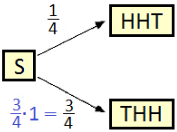
Game HHT-THH


S
H
HH
HHT
T
TH
THH
Rule 3
Rules 1 & 4
$\frac{\frac{1}{2}}{1-\frac{1}{2}} = 1$
$\frac{1}{2} \cdot 1 = \frac{1}{2}$
$\frac{1 \cdot \frac{1}{2}}{1 - 1 \cdot \frac{1}{2}} = 1$

Game HHT-THH

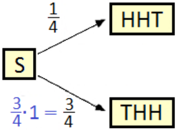

S
H
HH
HHT
T
TH
THH
Rule 3
Rule 1
Rules 1 & 4
$\frac{1}{2}\cdot\frac{1}{2}=\frac{1}{4}$
$\frac{1}{2}\cdot\frac{1}{2}=\frac{1}{4}$
$\frac{\frac{1}{2}}{1-\frac{1}{2}}=1$
$\frac{\frac{1}{2}}{1-\frac{1}{2}}=1$
$\frac{1}{2}\cdot 1=\frac{1}{2}$
$\frac{1\cdot\frac{1}{2}}{1-1\cdot\frac{1}{2}}=1$

Game HHT-THH



$$P(HHT) = \frac{1}{4}, \quad P(THH) = \frac{3}{4}$$

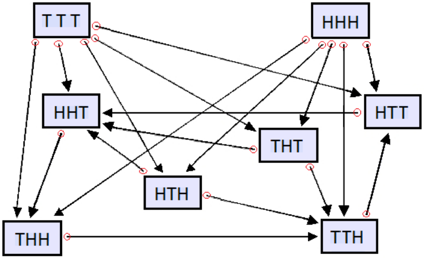
Game HHT-THH



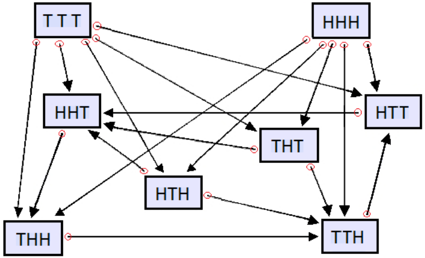
$$P(HHT) = \frac{1}{4}, \quad P(THH) = \frac{3}{4}$$

Among 28 classic Penney's games only 10 are fair!

Relations between 3-bit long sequences

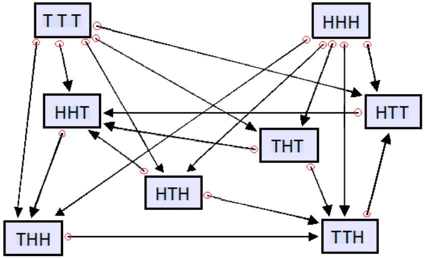


Relations between 3-bit long sequences



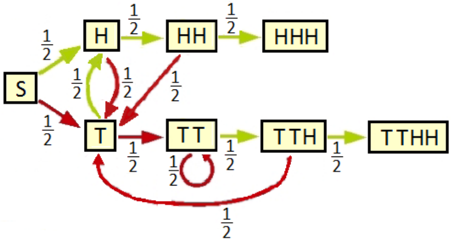
- There is no *the best* sequence in the classic Penney Game. Therefore priority of selection isn't a privilege in this game.

Relations between 3-bit long sequences

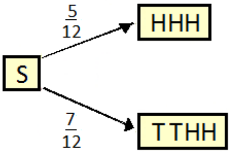


- There is no *the best* sequence in the classic Penney Game. Therefore priority of selection isn't a privilege in this game.
- The property of *being better than* in Penney's game isn't transitive.

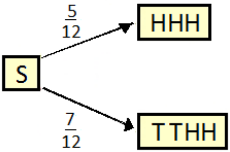
Games with sequences of different lengths - Game HHH-TTHH



Games with sequences of different lengths - Game HHH-TTHH



Games with sequences of different lengths - Game HHH-TTHH



$$P(TTHH) = \frac{7}{12} > \frac{1}{2}$$

Algorithm of calculating mean time of random walk on a stochastic graph with a nonempty boundary

Let (S, Q) be a stochastic graph with a nonempty boundary $B \subseteq S$ and T_j a random variable defined as the time of random walk begun in vertice $j \in S$ (and ending in the boundary B). Let us call:

- $E(T_j) = e_j$
- for $j \notin B$ we define K_j - the set of vertices attained directly from the vertice j $(k \in K_j \Leftrightarrow p_{jk} > 0)$

Algorithm of calculating mean time of random walk on a stochastic graph with a nonempty boundary

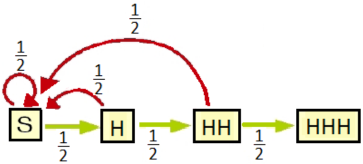
Let (S, Q) be a stochastic graph with a nonempty boundary $B \subseteq S$ and T_j a random variable defined as the time of random walk begun in vertice $j \in S$ (and ending in the boundary B). Let us call:

- $E(T_j) = e_j$
- for $j \notin B$ we define K_j - the set of vertices attained directly from the vertice j $(k \in K_j \Leftrightarrow p_{jk} > 0)$

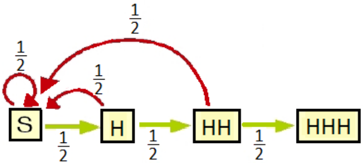
Then:

1. $e_j = 0 \quad$ for $j \in B$
2. $e_j = \sum_{k \in K_j} p_{jk} \cdot e_k + 1 \quad$ for $j \notin B$.

Mean time of waiting for sequence *HHH*

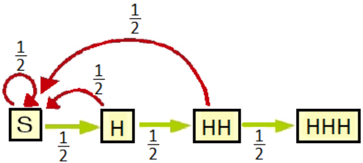


Mean time of waiting for sequence *HHH*



$$\begin{cases} e_S &= 1 + \frac{1}{2}\cdot e_S + \frac{1}{2}\cdot e_H \\ e_H &= 1 + \frac{1}{2}\cdot e_S + \frac{1}{2}\cdot e_{HH} \\ e_{HH} &= 1 + \frac{1}{2}\cdot e_S + \frac{1}{2}\cdot e_{HHH} \\ e_{HHH} &= 0 \end{cases} \qquad \begin{aligned} &\Rightarrow e_S = 14 \\ &\Rightarrow e_H = 12 \\ &\Rightarrow e_{HH} = 8 \\ &\Rightarrow e_{HHH} = 0 \end{aligned}$$

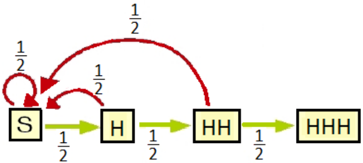
Mean time of waiting for sequence *HHH*



$$\begin{cases} e_S &= 1 + \frac{1}{2}\cdot e_S + \frac{1}{2}\cdot e_H \\ e_H &= 1 + \frac{1}{2}\cdot e_S + \frac{1}{2}\cdot e_{HH} \\ e_{HH} &= 1 + \frac{1}{2}\cdot e_S + \frac{1}{2}\cdot e_{HHH} \\ e_{HHH} &= 0 \end{cases} \qquad \begin{aligned} &\Rightarrow e_S = 14 \\ &\Rightarrow e_H = 12 \\ &\Rightarrow e_{HH} = 8 \\ &\Rightarrow e_{HHH} = 0 \end{aligned}$$

$\mathbb{E}T_{HHH} = 14, \qquad \mathbb{E}T_{TTHH} = 16$

Mean time of waiting for sequence *HHH*



$$\begin{cases} e_S & = 1 + \frac{1}{2} \cdot e_S + \frac{1}{2} \cdot e_H \\ e_H & = 1 + \frac{1}{2} \cdot e_S + \frac{1}{2} \cdot e_{HH} \\ e_{HH} & = 1 + \frac{1}{2} \cdot e_S + \frac{1}{2} \cdot e_{HHH} \\ e_{HHH} & = 0 \end{cases} \qquad \begin{aligned} &\Rightarrow e_S = 14 \\ &\Rightarrow e_H = 12 \\ &\Rightarrow e_{HH} = 8 \\ &\Rightarrow e_{HHH} = 0 \end{aligned}$$

$$\mathbb{E}T_{HHH} = 14, \qquad \mathbb{E}T_{TTHH} = 16$$

T_A - time of waiting until pattern A appears the first time

Notation

- A, B - arbitrary sequences of heads and tails, $|A| = k, |B| = m$, for $k, m > 0$;
- $A^{(i)}$ - first i elements of A, $A_{(i)}$ - last i elements of A, $i \leq k$;

Notation

- A, B - arbitrary sequences of heads and tails, $|A| = k, |B| = m$, for $k, m > 0$;
- $A^{(i)}$ - first i elements of A, $A_{(i)}$ - last i elements of A, $i \leq k$;
- $f(n)$ - the number of n-bit long sequences, which do not include pattern A, $n > 0$;
- $f_A(n)$ - the number of n-bit long sequences, which include pattern A at the end, and A doesn't appear before;

Notation

- A, B - arbitrary sequences of heads and tails, $|A| = k, |B| = m$, for $k, m > 0$;
- $A^{(i)}$ - first i elements of A, $A_{(i)}$ - last i elements of A, $i \leq k$;
- $f(n)$ - the number of n-bit long sequences, which do not include pattern A, $n > 0$;
- $f_A(n)$ - the number of n-bit long sequences, which include pattern A at the end, and A doesn't appear before;
- for $i = 1, 2, ..., min(k, m)$ we define

$$\delta_i(A, B) = \begin{cases} 2^i & \text{if } A_{(i)} = B^{(i)}, \\ 0 & \text{otherwise.} \end{cases} \qquad = 2^i \cdot [A_{(i)} = B^{(i)}]$$

Derivation of the formula for the mean time of waiting for a sequence

$$\left| \text{the set of } n\text{-bit long sequences without pattern } A \right| = \left| \text{the set of } (n+k)\text{-bit long sequences with pattern } A \text{ at the end, in which } A \text{ doesn't appear before } n\text{-th position} \right|$$

Derivation of the formula for the mean time of waiting for a sequence

$$\left|\text{the set of } n\text{-bit long sequences without pattern } A\right| = \left|\text{the set of } (n+k)\text{-bit long sequences with pattern } A \text{ at the end, in which } A \text{ doesn't appear before } n\text{-th position}\right|$$

$\nearrow \quad = f(n)$ $\qquad\qquad$ $\nearrow \quad \geq f_A(n+k)$

Derivation of the formula for the mean time of waiting for a sequence

$$\left| \text{the set of } n\text{-bit long sequences without pattern } A \right| = \left| \text{the set of } (n+k)\text{-bit long sequences with pattern } A \text{ at the end, in which } A \text{ doesn't appear before } n\text{-th position} \right|$$

$A = HTH, n = 4,$ $TTHT$

Derivation of the formula for the mean time of waiting for a sequence

$$\left| \text{the set of } n\text{-bit long sequences without pattern } A \right| = \left| \text{the set of } (n+k)\text{-bit long sequences with pattern } A \text{ at the end, in which } A \text{ doesn't appear before } n\text{-th position} \right|$$

$A = HTH, n = 4,$ $\quad TTHT \Rightarrow TT\underline{HT}\,\underline{H}TH$

Derivation of the formula for the mean time of waiting for a sequence

$$\left|\text{the set of } n\text{-bit long sequences without pattern } A\right| = \left|\text{the set of } (n+k)\text{-bit long sequences with pattern } A \text{ at the end, in which } A \text{ doesn't appear before } n\text{-th position}\right|$$

$A = HTH, n = 4, \qquad TTHT \Rightarrow TT\underline{HT}\,\underline{H}TH \qquad (A_{(1)} = A^{(1)})$

Derivation of the formula for the mean time of waiting for a sequence

$$\left| \text{the set of } n\text{-bit long sequences without pattern } A \right| = \left| \text{the set of } (n+k)\text{-bit long sequences with pattern } A \text{ at the end, in which } A \text{ doesn't appear before } n\text{-th position} \right|$$

$A = HTH, n = 4,$ $\quad TTHT \Rightarrow TT\underline{HT}\,\underline{H}TH$ $\quad (A_{(1)} = A^{(1)})$

$$f(n) = f_A(n+1) \cdot [A_{(1)} = A^{(1)}] + \cdots + f_A(n+k) \cdot [A_{(k)} = A^{(k)}]$$

Derivation of the formula for the mean time of waiting for a sequence

the set of n-bit long sequences without pattern A	=	the set of $(n+k)$-bit long sequences with pattern A at the end, in which A doesn't appear before n-th position

$A = HTH, n = 4, \qquad TTHT \Rightarrow TT\underline{HT}\ \underline{H}TH \qquad (A_{(1)} = A^{(1)})$

$$f(n) = f_A(n+1) \cdot [A_{(1)} = A^{(1)}] + \cdots + f_A(n+k) \cdot [A_{(k)} = A^{(k)}]$$

$$\frac{f(n)}{2^n} = \frac{f_A(n+1)}{2^{n+1}} \cdot \underbrace{2 \cdot [A_{(1)} = A^{(1)}]}_{\delta_1(A,A)} + \cdots + \frac{f_A(n+k)}{2^{n+k}} \cdot \underbrace{2^k \cdot [A_{(k)} = A^{(k)}]}_{\delta_k(A,A)}$$

Derivation of the formula for the mean time of waiting for a sequence

$$\left|\text{the set of } n\text{-bit long sequences without pattern } A\right| = \left|\text{the set of } (n+k)\text{-bit long sequences with pattern } A \text{ at the end, in which } A \text{ doesn't appear before } n\text{-th position}\right|$$

$$A = HTH, n = 4, \qquad TTHT \Rightarrow TT\underline{HT}\,\underline{H}TH \qquad (A_{(1)} = A^{(1)})$$

$$f(n) = f_A(n+1) \cdot [A_{(1)} = A^{(1)}] + \cdots + f_A(n+k) \cdot [A_{(k)} = A^{(k)}]$$

$$\frac{f(n)}{2^n} = \frac{f_A(n+1)}{2^{n+1}} \cdot \underbrace{2 \cdot [A_{(1)} = A^{(1)}]}_{\delta_1(A,A)} + \cdots + \frac{f_A(n+k)}{2^{n+k}} \cdot \underbrace{2^k \cdot [A_{(k)} = A^{(k)}]}_{\delta_k(A,A)}$$

T_A - time of waiting until pattern A appears the first time

Derivation of the formula for the mean time of waiting for a sequence

$$\left| \text{the set of } n\text{-bit long sequences without pattern } A \right| = \left| \text{the set of } (n+k)\text{-bit long sequences with pattern } A \text{ at the end, in which } A \text{ doesn't appear before } n\text{-th position} \right|$$

$A = HTH, n = 4,\quad TTHT \Rightarrow TT\underline{HT}\,\underline{H}TH \quad (A_{(1)} = A^{(1)})$

$$f(n) = f_A(n+1) \cdot [A_{(1)} = A^{(1)}] + \cdots + f_A(n+k) \cdot [A_{(k)} = A^{(k)}]$$

$$\frac{f(n)}{2^n} = \frac{f_A(n+1)}{2^{n+1}} \cdot \underbrace{2 \cdot [A_{(1)} = A^{(1)}]}_{\delta_1(A,A)} + \cdots + \frac{f_A(n+k)}{2^{n+k}} \cdot \underbrace{2^k \cdot [A_{(k)} = A^{(k)}]}_{\delta_k(A,A)}$$

T_A - time of waiting until pattern A appears the first time

$$P(T_A > n) = P(T_A = n+1) \cdot \delta_1(A,A) + \cdots + P(T_A = n+k) \cdot \delta_k(A,A)$$

Derivation of formula of mean time of waiting for the sequence...

$$P(T_A > n) = P(T_A = n+1)\cdot\delta_1(A,A) + \cdots + P(T_A = n+k)\cdot\delta_k(A,A)$$

Derivation of formula of mean time of waiting for the sequence...

$$P(T_A > n) = P(T_A = n+1)\cdot\delta_1(A,A) + \cdots + P(T_A = n+k)\cdot\delta_k(A,A)$$

$$\sum_{n=0}^{\infty} P(T_A > n) = \delta_1(A,A)\cdot\underbrace{\sum_{n=0}^{\infty} P(T_A = n+1)}_{1} + \cdots +$$

$$+\,\delta_k(A,A)\cdot\underbrace{\sum_{n=0}^{\infty} P(T_A = n+k)}_{1}$$

Derivation of formula of mean time of waiting for the sequence...

$$P(T_A > n) = P(T_A = n+1)\cdot\delta_1(A,A) + \cdots + P(T_A = n+k)\cdot\delta_k(A,A)$$

$$\sum_{n=0}^{\infty} P(T_A > n) = \delta_1(A,A)\cdot\underbrace{\sum_{n=0}^{\infty} P(T_A = n+1)}_{1} + \cdots + \\ + \delta_k(A,A)\cdot\underbrace{\sum_{n=0}^{\infty} P(T_A = n+k)}_{1}$$

$$\mathbb{E}T_A = \delta_1(A,A) + \cdots + \delta_k(A,A)$$

Derivation of formula of mean time of waiting for the sequence...

$$P(T_A > n) = P(T_A = n+1) \cdot \delta_1(A, A) + \cdots + P(T_A = n+k) \cdot \delta_k(A, A)$$

$$\sum_{n=0}^{\infty} P(T_A > n) = \delta_1(A, A) \cdot \underbrace{\sum_{n=0}^{\infty} P(T_A = n+1)}_{1} + \cdots + \\ + \delta_k(A, A) \cdot \underbrace{\sum_{n=0}^{\infty} P(T_A = n+k)}_{1}$$

$$\mathbb{E}T_A = \delta_1(A, A) + \cdots + \delta_k(A, A) \quad =: A : A$$

Conway's Formula:

$$\frac{p_B}{p_A} = \frac{A:A - A:B}{B:B - B:A}, \quad \text{where } A:B = \sum_{i=1}^{min(|A|,|B|)} \delta_i(A,B)$$

Conway's Formula:

$$\frac{p_B}{p_A} = \frac{A:A - A:B}{B:B - B:A}, \quad \text{where } A:B = \sum_{i=1}^{min(|A|,|B|)} \delta_i(A,B)$$

- for any $n > 0$ longer sequence $\underbrace{T\ldots T}_{n}\underbrace{H\ldots H}_{n}$ is *better* than shorter one - $\underbrace{H\ldots H}_{2n-1}$

Conway's Formula:

$$\frac{p_B}{p_A} = \frac{A:A - A:B}{B:B - B:A}, \quad \text{where } A:B = \sum_{i=1}^{min(|A|,|B|)} \delta_i(A,B)$$

- for any $n > 0$ longer sequence $\underbrace{T...T}_{n}\underbrace{H...H}_{n}$ is *better* than shorter one - $\underbrace{H...H}_{2n-1}$
- for sequences of any length there is no *best sequence*: we know even more - for arbitrary sequence $a_1 \dots a_m$ one of sequences of the form $ba_1 \dots a_{m-1}$, where b is H or T, is the best in the game with $a_1 \dots a_m$;

Conway's Formula:

$$\frac{p_B}{p_A} = \frac{A:A - A:B}{B:B - B:A}, \quad \text{where } A:B = \sum_{i=1}^{min(|A|,|B|)} \delta_i(A,B)$$

- for any $n > 0$ longer sequence $\underbrace{T...T}_{n}\underbrace{H...H}_{n}$ is *better* than shorter one - $\underbrace{H...H}_{2n-1}$
- for sequences of any length there is no *best sequence*: we know even more - for arbitrary sequence $a_1 \ldots a_m$ one of sequences of the form $ba_1 \ldots a_{m-1}$, where b is H or T, is the best in the game with $a_1 \ldots a_m$;
- Penney's game with sequences of any length $n > 4$ isn't transitive.

Why Penney's game is interesting for us?

- The simplicity of the game and the its surprising results make it an interesting tool for mathematics popularization.

Why Penney's game is interesting for us?

- The simplicity of the game and the its surprising results make it an interesting tool for mathematics popularization.
- We can prove an analogous version of the Conway's formula for a game with a countable set of results obtained by tossing an arbitrary die instead of a symmetric coin. We can obtain set of linear equations equivalent to Conway's Formula for games with several players.

Why Penney's game is interesting for us?

- The simplicity of the game and the its surprising results make it an interesting tool for mathematics popularization.
- We can prove an analogous version of the Conway's formula for a game with a countable set of results obtained by tossing an arbitrary die instead of a symmetric coin. We can obtain set of linear equations equivalent to Conway's Formula for games with several players.
- This general theory will perhaps find some applications i.e. in:
  - mathematical modeling of gene mutations;
  - game theory and its applications i.e. on stock exchange.

Thank You for your attention!