Coin racing and waiting-time paradoxes: why fair coins are exceptional

Independent pattern races, stochastic dominance, and non-transitive tournaments

Søren Riis

Queen Mary University of London
(joint work with Mike Paterson)

Paper: arXiv:2601.16580

arXiv:2601.16580

A bus-stop puzzle

You arrive at a bus stop.

The sign says: “On average, a bus comes every 16 minutes.”

Question: what is your *expected waiting time*?

- 8 minutes? (“half of 16”)
- 16 minutes? (“random arrivals”)
- **It depends**? (regular vs clumpy)

A bus-stop puzzle

You arrive at a bus stop.

The sign says: "On average, a bus comes every 16 minutes."

Question: what is your *expected waiting time*?

- 8 minutes? ("half of 16")
- 16 minutes? ("random arrivals")
- **It depends**? (regular vs clumpy)

The answer

It depends on regularity. Perfectly regular buses give 8 minutes; memoryless arrivals (Poisson / geometric) give 16 minutes; *clumping* pushes the wait back up.

Example of clumping: buses come in pairs, 1 minute apart, then a 31-minute gap (mean gap 16). Long empty stretches dominate what you experience.

Same frequency, different first-hit times

On a **fair coin**, consider these length-4 patterns:

$$\text{HHTT}, \quad \text{HTTH}, \quad \text{HTHT}, \quad \text{HHHH}.$$

Fact (frequency): each length-4 pattern has probability $2^{-4} = 1/16$, so in a long run you see each pattern about once every 16 tosses.

But we ask a different question: starting from scratch, what is the mean time $\mathbb{E}[\tau_T]$ until the *first* occurrence?

Coin-toss buses: why "every 16" does *not* mean "wait 16"

Think of each pattern as a **bus line**.

Watch one coin stream $X_1X_2X_3\ldots$. A bus of line T arrives at time n if

$$X_{n-3}X_{n-2}X_{n-1}X_n = T.$$

Why the wait can be long:

- buses can **clump** because patterns can *overlap*,

Coin-toss buses: why "every 16" does *not* mean "wait 16"

Think of each pattern as a **bus line**.

Watch one coin stream $X_1X_2X_3\ldots$. A bus of line T arrives at time n if

$$X_{n-3}X_{n-2}X_{n-1}X_n = T.$$

Why the wait can be long:

- buses can **clump** because patterns can *overlap*,
- clumps create **long empty stretches** elsewhere,

Coin-toss buses: why "every 16" does *not* mean "wait 16"

Think of each pattern as a **bus line**.

Watch one coin stream $X_1X_2X_3\ldots$. A bus of line T arrives at time n if

$$X_{n-3}X_{n-2}X_{n-1}X_n = T.$$

Why the wait can be long:

- buses can **clump** because patterns can *overlap*,
- clumps create **long empty stretches** elsewhere,
- so the first bus can take *longer* than the mean gap.

Coin-toss buses: why "every 16" does *not* mean "wait 16"

Think of each pattern as a **bus line**.

Watch one coin stream $X_1 X_2 X_3 \ldots$. A bus of line T arrives at time n if

$$X_{n-3} X_{n-2} X_{n-1} X_n = T.$$

Why the wait can be long:

- buses can **clump** because patterns can *overlap*,
- clumps create **long empty stretches** elsewhere,
- so the first bus can take *longer* than the mean gap.

Example (clumping)

`...H H H H H ...`
contains *two* `HHHH` buses one toss apart.

Rule of thumb

If the "bus indicator" at each step had marginal probability p *and* were memoryless, the mean wait would be $1/p$. Overlaps break memorylessness, and can only increase the mean first-hit time.

Borders (prefix–suffix overlaps) drive the waiting-time distribution

A **border** of a word T is a nonempty string that is both a prefix and a suffix of T.

Pattern	**Borders**	**Border lengths**
HHTT	HHTT	$\{4\}$
HTTH	H, HTTH	$\{1,4\}$
HTHT	HT, HTHT	$\{2,4\}$
HHHH	H, HH, HHH, HHHH	$\{1,2,3,4\}$

Key point: the set of border lengths (always including $|T|$) determines the entire waiting-time distribution, not just the mean.

The border-sum formula (fair coin)

For a fair coin:

$$\mathbb{E}[\tau_T] = \sum_{\ell \in \mathcal{B}(T)} 2^{\ell}.$$

Interpretation:

- the leading term $2^{|T|}$ is the "frequency" term $1/\mathbb{P}(T)$;
- each *proper* border contributes an extra startup penalty.

The warm-up, resolved

Pattern	$\mathcal{B}(T)$	$\mathbb{E}[\tau_T]$
HHTT	$\{4\}$	16
HTTT	$\{4\}$	16
HTTH	$\{1,4\}$	$2+16=18$
HTHT	$\{2,4\}$	$4+16=20$
HHHH	$\{1,2,3,4\}$	$2+4+8+16=30$

The warm-up, resolved

Pattern	$\mathcal{B}(T)$	$\mathbb{E}[\tau_T]$
HHTT	$\{4\}$	16
HTTT	$\{4\}$	16
HTTH	$\{1,4\}$	$2+16=18$
HTHT	$\{2,4\}$	$4+16=20$
HHHH	$\{1,2,3,4\}$	$2+4+8+16=30$

Unbordered patterns win the mean race; **runs** lose badly.

A striking identity

On a fair coin:

$$\mathbb{E}[\tau_{\mathrm{HHHH}}] = 2 + 4 + 8 + 16 = 30.$$

A striking identity

On a fair coin:

$$\mathbb{E}[\tau_{\mathrm{HHHH}}] = 2 + 4 + 8 + 16 = 30.$$

But also

$$\mathbb{E}[\tau_{\mathrm{HHHT}}] = 16, \quad \mathbb{E}[\tau_{\mathrm{HHT}}] = 8, \quad \mathbb{E}[\tau_{\mathrm{HT}}] = 4, \quad \mathbb{E}[\tau_{\mathrm{T}}] = 2.$$

So

$$\mathbb{E}[\tau_{\mathrm{HHHH}}] = \mathbb{E}[\tau_{\mathrm{HHHT}}] + \mathbb{E}[\tau_{\mathrm{HHT}}] + \mathbb{E}[\tau_{\mathrm{HT}}] + \mathbb{E}[\tau_{\mathrm{T}}].$$

Intuition: the self-overlaps in HHHH create "checkpoints" at lengths 1,2,3,4.

Pattern matching as a Markov chain (prefix automaton)

For a fixed pattern $T = t_1 \cdots t_L$, build an automaton that tracks the length of the *longest matched prefix*.

States: one for each prefix length $0, 1, \ldots, L$.

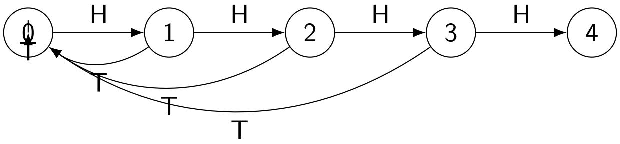
- State 0 is the empty prefix (no progress).
- State i $(1 \leq i < L)$ means the last i tosses equal $t_1 \cdots t_i$.
- State L is the full pattern—the absorbing state.

State count: L transient states (plus one absorbing).

Where borders enter: when a mismatch occurs at state i, the chain *falls back* to the largest prefix length that is also a suffix of the current history. This fallback structure is encoded by the border lengths.

Example: the run H^4 has a very simple chain

For $T = \mathrm{HHHH}$, every prefix is a border, so the automaton is a straight line:



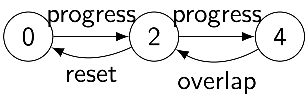
This immediately yields the recursion that solves to $\mathbb{E}[\tau_{\mathrm{H}^4}] = 2 + 4 + 8 + 16$.

Borders as a "fallback graph" (intuition for overlap)

The prefix automaton can be viewed as having **failure links** that jump to shorter matched prefixes. Those jumps are governed by *borders* (prefix–suffix overlaps).

For intuition, it is helpful to draw a **border graph** whose nodes are the border lengths $\mathcal{B}(T) \cup \{0\}$ and whose directed edges indicate typical fallback targets.

Example: $T = \mathrm{HTHT}$ **(borders** $\{2, 4\}$**).**
After a hit, we immediately keep a length-2 match (the overlap HT), so hits *clump*.



Unbordered patterns: if $\mathcal{B}(T) = \{L\}$ only, the border graph is essentially a reset, so the mean collapses to $\mathbb{E}[\tau_T] = 1/\mathbb{P}(T)$.

A betting puzzle

You and a friend are at a bus stop, each waiting for a different bus line.
You bet on whose bus arrives first.

Question: Can your expected waiting time be *longer* than your friend's, yet you still have the better odds of winning the bet?

A betting puzzle

You and a friend are at a bus stop, each waiting for a different bus line.
You bet on whose bus arrives first.

Question: Can your expected waiting time be *longer* than your friend's, yet you still have the better odds of winning the bet?

Answer: Yes.

The mean waiting time and the probability of winning a race measure different things:

- The **mean** averages over the whole distribution—it's sensitive to rare but very long waits.
- **Winning a race** depends on having more probability mass *early*.

A distribution can have good early odds (wins races) but a heavy tail (inflates the mean).

Mean waiting time vs winning a race

It is possible for one pattern to have a *larger* mean waiting time, yet still have *better* odds of beating another pattern in a head-to-head race.

Two notions of "faster":

- **Mean** $\mathbb{E}[\tau]$: averages over the whole distribution (sensitive to rare but very long waits).
- **Race odds** $\mathbb{P}(\tau_A < \tau_B)$: depends on having more probability mass *early*.

So a pattern can win races because it is more likely to hit early, even if its tail is heavy enough to inflate the mean.

Two patterns racing: product chain (why it gets big)

For a head-to-head race between patterns A (length L_A) and B (length L_B):

- build the prefix automaton for A (size L_A), and for B (size L_B);
- take their **product** to track both players simultaneously.

State count: at most $L_A L_B$ transient states, plus absorbing states "A wins", "B wins", and "tie".

Consequences:

- probabilities are solutions to a finite linear system;
- with a common bias p, each win probability is a **rational function of** p.

Generating functions for waiting times

For a pattern T, let τ_T be the first-hit time and set

$$a_{T,n} := \mathbb{P}(\tau_T = n), \qquad S_T(n) := \mathbb{P}(\tau_T > n).$$

Probability generating function (pgf):

$$A_T(x) = \sum_{n\geq 1} a_{T,n}\, x^n.$$

Survival generating function:

$$\widetilde{B}_T(x) = \sum_{n\geq 1} S_T(n)\, x^n.$$

Tail/pgf relation:

$$\widetilde{B}_T(x) = \frac{x - A_T(x)}{1-x}.$$

Two payoffs: $A_T(1) = 1,$ and $\mathbb{E}[\tau_T] = A'_T(1).$

Key fact: $A_T(x)$ is a *rational function* (finite prefix automaton $\Rightarrow$ finite absorbing Markov chain).

How borders shape $A_T(x)$ (fair die / fair coin)

Let T have length L, and let $\mathcal{B}(T)$ be its set of border lengths (including L).

Border polynomial:

$$H_T(s) = \sum_{\ell \in \mathcal{B}(T)} s^{\ell},$$

where s is the alphabet size ($s = 2$ for a coin).

For a **fair s-sided die**, the whole hitting-time law can be encoded by the same border data:

$$\boxed{A_T(x) = \frac{x^L}{x^L + (1-x) \sum_{\ell \in \mathcal{B}(T)} s^{\ell} \, x^{L-\ell}}} \qquad \text{(fair source)}.$$

Differentiating at $x = 1$ gives back the border-sum mean formula:

$$\mathbb{E}[\tau_T] = A'_T(1) = \sum_{\ell \in \mathcal{B}(T)} s^{\ell} \quad \left(\text{so for a fair coin, } \mathbb{E}[\tau_T] = \sum_{\ell \in \mathcal{B}(T)} 2^{\ell}\right).$$

Example: HHTT vs HHHH

Both patterns have probability $2^{-4} = 1/16$ (so the *mean recurrence gap* is 16), but their *first-hit* laws differ because of borders.

HHTT (unbordered)
$\mathcal{B}(\mathtt{HHTT}) = \{4\}$

$$A_{\mathtt{HHTT}}(x) = \frac{x^4}{x^4 + 16(1-x)}.$$

$$\mathbb{E}[\tau_{\mathtt{HHTT}}] = 16.$$

HHHH (maximal overlap)
$\mathcal{B}(\mathtt{HHHH}) = \{1, 2, 3, 4\}$

$$A_{\mathtt{HHHH}}(x) = \frac{x^4}{x^4 + (1-x)(16 + 8x + 4x^2 + 2x^3)}.$$

$$\mathbb{E}[\tau_{\mathtt{HHHH}}] = 30.$$

Interpretation: extra borders $\Rightarrow$ more overlap $\Rightarrow$ more clumping $\Rightarrow$ heavier tail (and larger mean).

From single patterns to races (Hadamard products)

For independent streams, the strict win probability decomposes time-by-time:

$$\mathbb{P}(\tau_A < \tau_B) = \sum_{n \geq 1} \mathbb{P}(\tau_A = n)\, \mathbb{P}(\tau_B > n) = \sum_{n \geq 1} a_{A,n}\, S_B(n).$$

In generating-function form this is a coefficientwise pairing. Define the **Hadamard product**

$$(F \odot G)(x) := \sum_{n \geq 0} f_n g_n x^n \quad \text{for } F(x) = \sum f_n x^n,\ \ G(x) = \sum g_n x^n.$$

With $\widetilde{B}_B(x) = \sum_{n \geq 1} S_B(n) x^n = \frac{x - A_B(x)}{1-x}$,

$$\boxed{\mathbb{P}(\tau_A < \tau_B) = \lim_{x \uparrow 1} (A_A \odot \widetilde{B}_B)(x).}$$

Random tie-break adds $\frac{1}{2} \lim_{x \uparrow 1} (A_A \odot A_B)(x)$.

Key point: race odds can be computed using only the *single-pattern* rational functions $A_T(x)$.

Stochastic dominance = "wins at every checkpoint"

For a pattern T, let τ_T be its first-hit time.

Stochastic dominance:

$$T_1 \succ_{\mathrm{st}} T_2 \quad \iff \quad \mathbb{P}(\tau_{T_1} \le m) \ge \mathbb{P}(\tau_{T_2} \le m) \text{ for all } m \ge 0, \text{ and strict for some } m.$$

Intuition: T_1 is *ahead* of T_2 at **every checkpoint**—it has a higher probability of having already appeared. Since $\mathbb{E}[\tau] = \sum_{m \ge 0} \mathbb{P}(\tau > m)$, dominance forces mean ordering: $T_1 \succ_{\mathrm{st}} T_2 \Rightarrow \mathbb{E}[\tau_{T_1}] < \mathbb{E}[\tau_{T_2}]$.

Fair die theorem

Theorem (fairness = total order). For a fair s-sided die, stochastic dominance totally pre-orders *all strings*. Moreover, under fairness:

$$\tau_{T_1} \succ_{\mathrm{st}} \tau_{T_2} \quad \Longleftrightarrow \quad \mathbb{E}[\tau_{T_1}] \geq \mathbb{E}[\tau_{T_2}].$$

Why:

- in the fair case, the border polynomial $H_T(s) = \sum_{\ell \in \mathcal{B}(T)} s^\ell$ uniquely determines the waiting-time law;
- comparing border-indicator vectors lexicographically yields a fixed sign across all checkpoints.

Biased coins: total comparability fails

Let $\mathbb{P}(\mathrm{H}) = p$, $\mathbb{P}(\mathrm{T}) = q = 1 - p$. Assume $p > q$ (the other case is symmetric).

Pick $n \geq 2$ such that $p^n < q$ and compare

$$A = \mathrm{H}^n, \qquad B = \mathrm{H}^{n-1}\mathrm{T}.$$

The sprint (time $m = n$):

$$\mathbb{P}(\tau_A = n) = p^n \ > \ p^{n-1}q = \mathbb{P}(\tau_B = n) \quad \Rightarrow \quad \mathbb{P}(\tau_A > n) < \mathbb{P}(\tau_B > n).$$

So A cannot dominate B.

Biased coins: total comparability fails

Let $\mathbb{P}(\mathrm{H}) = p$, $\mathbb{P}(\mathrm{T}) = q = 1 - p$. Assume $p > q$ (the other case is symmetric).

Pick $n \geq 2$ such that $p^n < q$ and compare

$$A = \mathrm{H}^n, \qquad B = \mathrm{H}^{n-1}\mathrm{T}.$$

The sprint (time $m = n$):

$$\mathbb{P}(\tau_A = n) = p^n \; > \; p^{n-1}q = \mathbb{P}(\tau_B = n) \quad \Rightarrow \quad \mathbb{P}(\tau_A > n) < \mathbb{P}(\tau_B > n).$$

So A cannot dominate B.

The marathon (mean):

$$\mathbb{E}[\tau_A] = \sum_{j=1}^{n} p^{-j} = \frac{p^{-n} - 1}{q} \; > \; \frac{1}{p^{n-1}q} = \mathbb{E}[\tau_B]$$

(the inequality is exactly $p^n < q$). Dominance would force the opposite mean order, so B cannot dominate A either.

Biased coins: total comparability fails

Let $\mathbb{P}(\mathrm{H}) = p$, $\mathbb{P}(\mathrm{T}) = q = 1 - p$. Assume $p > q$ (the other case is symmetric).

Pick $n \geq 2$ such that $p^n < q$ and compare

$$A = \mathrm{H}^n, \qquad B = \mathrm{H}^{n-1}\mathrm{T}.$$

The sprint (time $m = n$):

$$\mathbb{P}(\tau_A = n) = p^n \; > \; p^{n-1}q = \mathbb{P}(\tau_B = n) \quad \Rightarrow \quad \mathbb{P}(\tau_A > n) < \mathbb{P}(\tau_B > n).$$

So A cannot dominate B.

The marathon (mean):

$$\mathbb{E}[\tau_A] = \sum_{j=1}^{n} p^{-j} = \frac{p^{-n} - 1}{q} \; > \; \frac{1}{p^{n-1}q} = \mathbb{E}[\tau_B]$$

(the inequality is exactly $p^n < q$). Dominance would force the opposite mean order, so B cannot dominate A either.

Conclusion: τ_A and τ_B are incomparable under stochastic dominance when $p \neq \frac{1}{2}$.

The fairness dichotomy

Theorem. Let $p \neq \frac{1}{2}$ and choose $n \geq 2$ such that $p^n < 1 - p$. Then τ_A and τ_B (for $A = \mathrm{H}^n$, $B = \mathrm{H}^{n-1}\mathrm{T}$) are incomparable under stochastic dominance.

The fairness dichotomy

Theorem. Let $p \neq \frac{1}{2}$ and choose $n \geq 2$ such that $p^n < 1 - p$. Then τ_A and τ_B (for $A = \mathrm{H}^n$, $B = \mathrm{H}^{n-1}\mathrm{T}$) are incomparable under stochastic dominance.

Corollary. A coin is fair if and only if stochastic dominance totally orders all patterns.

Equivalently: a coin is biased if and only if there exist patterns with incomparable waiting times under stochastic dominance.

Win probability and the advantage function

Two independent streams; random tie-break:

$$W^{\mathrm{rtb}}(A > B; p) = \mathbb{P}(\tau_A < \tau_B) + \tfrac{1}{2}\mathbb{P}(\tau_A = \tau_B).$$

Define the **advantage**:

$$g_{A,B}(p) := W^{\mathrm{rtb}}(A > B; p) - \tfrac{1}{2}.$$

Then $A \to B$ iff $g_{A,B}(p) > 0$.

Markov-chain route: why $g_{A,B}(p)$ is rational

The joint prefix automaton for (A, B) is a finite absorbing Markov chain. Absorption probabilities are entries of

$$\Pi(p) = (I - Q(p))^{-1} R(p),$$

so for common bias p each entry is a rational function in p.

Hence

$$g_{A,B}(p) = \frac{N_{A,B}(p)}{D_{A,B}(p)} \quad \text{with } N_{A,B}, D_{A,B} \in \mathbb{Z}[p].$$

Example (length 2): $A = \mathrm{HH}$ vs $B = \mathrm{HT}$. The threshold $g_{A,B}(p) = 0$ is the sextic

$$p^6 - 3p^5 + 2p^4 + p^2 + p - 1 = 0 \qquad (p_* \approx 0.586648\ldots).$$

Reversal: mean vs win probability (HH vs HT)

For HH vs HT with random tie-break:

- win probability crosses $1/2$ at $p_* \approx 0.586648$;
- means cross at $p_\varphi = (\sqrt{5} - 1)/2 \approx 0.618034$.

So on the window $p \in (p_*, p_\varphi)$ we have a genuine reversal:

$$W^{\mathrm{rtb}}(\mathrm{HH} > \mathrm{HT}; p) > \tfrac{1}{2} \quad \text{but} \quad \mathbb{E}[\tau_{\mathrm{HH}}] > \mathbb{E}[\tau_{\mathrm{HT}}].$$

Reversal: mean vs win probability (HH vs HT)

For HH vs HT with random tie-break:

- win probability crosses $1/2$ at $p_* \approx 0.586648$;
- means cross at $p_\varphi = (\sqrt{5} - 1)/2 \approx 0.618034$.

So on the window $p \in (p_*, p_\varphi)$ we have a genuine reversal:

$$W^{\mathrm{rtb}}(\mathrm{HH} > \mathrm{HT}; p) > \tfrac{1}{2} \quad \text{but} \quad \mathbb{E}[\tau_{\mathrm{HH}}] > \mathbb{E}[\tau_{\mathrm{HT}}].$$

The moral: “smaller mean waiting time” does *not* imply “better chance to win a head-to-head race” once the coin is biased.

A non-transitive bus puzzle

Three friends A, B, C are at a bus stop, each waiting for a different bus line. They bet pairwise on whose bus arrives first. Can the winner cycle?

A non-transitive bus puzzle

Three friends A, B, C are at a bus stop, each waiting for a different bus line. They bet pairwise on whose bus arrives first. Can the winner cycle?

Example: Suppose the buses run on a regular hourly schedule:

- Bus A arrives on the hour (:00)
- Bus B arrives at :20
- Bus C arrives at :40

If the friends arrive at a *uniformly random* time within the hour:

- A beats B: arrival in (:20, :00] $\Rightarrow$ next A comes before next B (40 min)
- B beats C: arrival in (:40, :20] $\Rightarrow$ next B comes before next C (40 min)
- C beats A: arrival in (:00, :40] $\Rightarrow$ next C comes before next A (40 min)

Each pairwise race is won with probability $\frac{2}{3}$, forming a cycle: $A > B > C > A$.

Paradoxical triples (win tournament cycles)

Fix a tie convention (we use random tie-break). For patterns A, B, C, define the directed edge

$$A \to B \iff W^{\mathrm{rtb}}(A > B; p) > \tfrac{1}{2}.$$

A **paradoxical triple** is a directed 3-cycle:

$$A \to B, \quad B \to C, \quad C \to A.$$

Paradoxical triples (win tournament cycles)

Fix a tie convention (we use random tie-break). For patterns A, B, C, define the directed edge

$$A \to B \quad \iff \quad W^{\mathrm{rtb}}(A > B; p) > \tfrac{1}{2}.$$

A **paradoxical triple** is a directed 3-cycle:

$$A \to B, \quad B \to C, \quad C \to A.$$

Key point: cycles are impossible under fairness for stochastic dominance, but they *do* occur under bias for the win tournament.

Common bias: the minimal $(2, 5, 5)$ cycle

Three players, *same* biased coin with $\mathbb{P}(\mathrm{H}) = p$.

Smallest common-bias cycle (up to symmetry):

$$T_1 = \mathrm{TT} \quad \text{(length 2)}, \qquad T_2 = \mathrm{HHHHT}, \qquad T_3 = \mathrm{HHHHH}.$$

Common bias: the minimal $(2, 5, 5)$ cycle

Three players, *same* biased coin with $\mathbb{P}(\mathrm{H}) = p$.

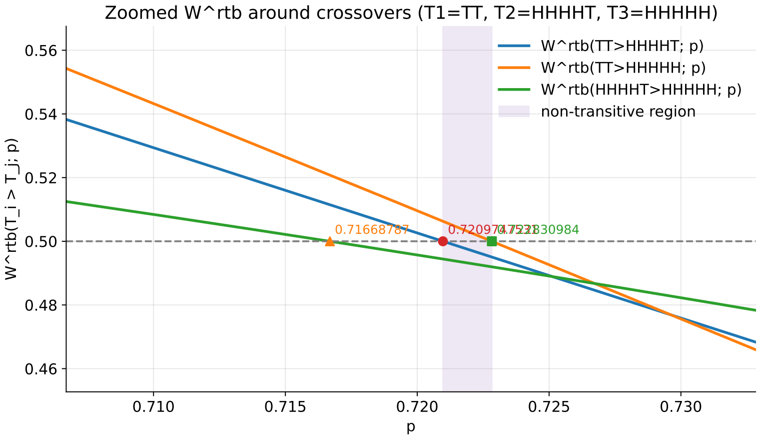
Smallest common-bias cycle (up to symmetry):

$$T_1 = \mathrm{TT} \quad \text{(length 2)}, \qquad T_2 = \mathrm{HHHHT}, \qquad T_3 = \mathrm{HHHHH}.$$

It forms a strict 3-cycle on the narrow window

$$p \in (0.720974753\ldots,\ 0.722830984\ldots).$$

Visual: the $(2, 5, 5)$ window



Exact classification (common bias, max length $L \leq 8$)

We performed an exhaustive certified search for common-bias 3-cycles with $\max(|A|, |B|, |C|) \leq 8$.

Result:

- exactly 16 cycle windows for $p \geq 1/2$ up to equivalence;
- by complement symmetry ($p \mapsto 1 - p$) this yields **32** windows in $(0, 1)$;
- none of the certified common-bias cycles uses three equal lengths (all have unequal length profiles).

Next two slides list the 16 windows for $p \geq 1/2$.

Complete classification of common-bias cycle windows for $L \leq 8$

Proposition. There are exactly 32 common-bias 3-cycle windows with $\max(|A|, |B|, |C|) \leq 8$: 16 for $p > 1/2$, and 16 mirrors for $p < 1/2$ (via $p \mapsto 1 - p$).

#	lengths	(A, B, C) (binary, $0 = \mathrm{H}, 1 = \mathrm{T}$)	p-window (approx.)
1	(2,5,5)	11, 00001, 00000	(0.72097, 0.72283)
2	(6,6,3)	000000, 000001, 011	(0.74274, 0.74448)
3	(6,6,3)	000000, 000010, 101	(0.72936, 0.73069)
4	(6,6,4)	000000, 000100, 0110	(0.70653, 0.70678)
5	(6,6,5)	000000, 001100, 10101	(0.59244, 0.59247)
6	(6,5,6)	000111, 11111, 100001	(0.50739, 0.50740)
7	(7,7,4)	0000000, 0000100, 1001	(0.73539, 0.73621)
8	(7,5,6)	0000100, 01101, 010001	(0.62149, 0.62150)

Continued on next slide (entries 9–16).

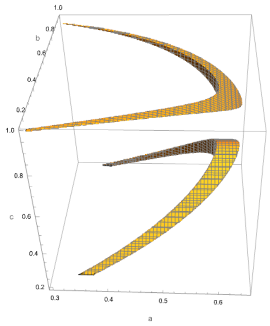
Complete classification of common-bias cycle windows for $L \leq 8$ (continued)

#	lengths	(A, B, C) (binary, $0 = \mathrm{H}, 1 = \mathrm{T}$)	p-window (approx.)
9	(7,3,5)	0000100, 111, 01010	(0.67205, 0.67251)
10	(8,8,3)	00000000, 00000001, 101	(0.78023, 0.78175)
11	(8,8,4)	00000000, 00000010, 0011	(0.77126, 0.77219)
12	(8,8,4)	00000000, 00000100, 0101	(0.75959, 0.76019)
13	(8,4,7)	00000010, 1111, 0000011	(0.62374, 0.62376)
14	(8,5,7)	00000010, 11011, 0001100	(0.61477, 0.61477)
15	(8,6,7)	00000100, 011001, 0100001	(0.62025, 0.62025)
16	(8,4,6)	00001000, 0111, 001100	(0.69639, 0.69650)

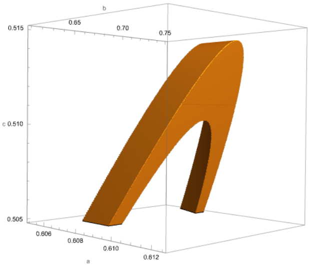
The 16 mirror windows for $p < 1/2$ are $(1 - p_{\mathrm{high}}, 1 - p_{\mathrm{low}})$.

Unequal biases: "McDonald's arches" (cycle volume)

For three fixed length-3 patterns HHH, HTH, HHT and independent coins with possibly different biases (p_A, p_B, p_C), the set of parameters yielding a 3-cycle forms a striking 3D region.



Zoomed view of one arch



Algebraic definition of the arch region (Sage / exact arithmetic)

Fix patterns

$$A = \mathrm{HHH}, \qquad B = \mathrm{HTH}, \qquad C = \mathrm{HHT},$$

and let players use *independent* coins with biases $p_A, p_B, p_C \in (0,1)$.

With random tie-break, define pairwise advantages

$$g_{A,B}(p_A, p_B) := W^{\mathrm{rtb}}(A > B; p_A, p_B) - \tfrac{1}{2}, \quad \text{etc.}$$

Sage produces each g as a rational function. After clearing denominators (and verifying the common denominator has fixed sign on $(0,1)^2$), each comparison reduces to the sign of a bivariate integer polynomial:

$$\begin{aligned} A \to B &\iff G_{AB}(p_A, p_B) > 0, \\ B \to C &\iff G_{BC}(p_B, p_C) > 0, \\ C \to A &\iff G_{CA}(p_C, p_A) > 0. \end{aligned}$$

Semi-algebraic description of the non-transitive region

$$\mathcal{R} = \{(p_A, p_B, p_C) \in (0,1)^3 : \; G_{AB} > 0, \; G_{BC} > 0, \; G_{CA} > 0\}.$$

In this example (Sage output) $\deg(G_{AB}) = \deg(G_{BC}) = \deg(G_{CA}) = 17$.

One defining polynomial (degree 17): $G_{AB}(p_A, p_B)$

$$\begin{aligned} G_{AB}(p_A,p_B) = {} & p_A^9p_B^8 - p_A^8p_B^9 - 5p_A^9p_B^7 + 3p_A^8p_B^8 + 2p_A^7p_B^9 + 10p_A^9p_B^6 \\ & - 8p_A^7p_B^8 - p_A^6p_B^9 - 10p_A^9p_B^5 - 10p_A^8p_B^6 + 11p_A^7p_B^7 + 3p_A^6p_B^8 + 5p_A^9p_B^4 \\ & + 15p_A^8p_B^5 - 4p_A^7p_B^6 - 3p_A^6p_B^7 + p_A^5p_B^8 - p_A^9p_B^3 - 9p_A^8p_B^4 - 5p_A^7p_B^5 \\ & + 2p_A^6p_B^6 - 2p_A^5p_B^7 + 2p_A^8p_B^3 + 7p_A^7p_B^4 - 2p_A^5p_B^6 - p_A^4p_B^7 - 4p_A^7p_B^3 \\ & - 5p_A^6p_B^4 + 6p_A^5p_B^5 + 2p_A^4p_B^6 + p_A^7p_B^2 + 6p_A^6p_B^3 - 2p_A^5p_B^4 - 2p_A^4p_B^5 \\ & + 2p_A^3p_B^6 - 2p_A^6p_B^2 - p_A^5p_B^3 + 2p_A^4p_B^4 - 3p_A^3p_B^5 - p_A^5p_B^2 - p_A^2p_B^5 \\ & + p_A^5p_B - p_A^4p_B^2 + p_A^3p_B^3 + 2p_A^2p_B^4 + p_A^3p_B^2 - 2p_A^2p_B^3 - p_A^3p_B + p_A^2p_B^2 \\ & - p_Ap_B^3 - p_A^3 + p_Ap_B^2 - p_B^3 + p_B^2. \end{aligned}$$

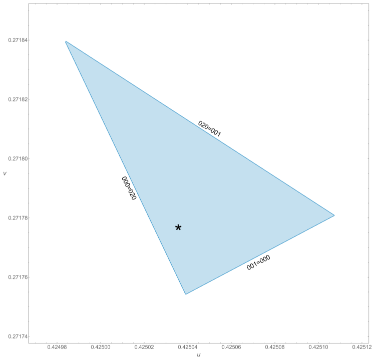
(Analogous degree-17 polynomials $G_{BC}(p_B, p_C)$ and $G_{CA}(p_C, p_A)$ decide the other two edges.)

A biased 3-die cycle region

Example on a biased 3-sided die (faces $0, 1, 2$):

$$000 > 020 > 001 > 000$$

on an open region of the simplex (p_0, p_1, p_2).



Looks like a triangle, but it isn't (semi-algebraic region)

In the 3-sided die example, write the simplex parameters as

$$(p_0, p_1, p_2) = (u, v, 1 - u - v), \qquad u > 0,\ v > 0,\ u + v < 1.$$

For patterns $A = 000,\ B = 020,\ C = 001$, the strict cycle

$$000 > 020 > 001 > 000$$

holds on a region $\mathcal{R}$ of the simplex that can be described *exactly* by polynomial inequalities:

$$\mathcal{R} = \Big\{(u, v) \in \mathbb{R}^2 :\ u > 0,\ v > 0,\ u+v < 1,\ -P_{000,020}(u, v) > 0,\ P_{020,001}(u, v) > 0,\ P_{001,000}(u, v) >$$

- Visually this region can look almost "triangular", but each boundary is an *algebraic curve* $P(u, v) = 0$ (degrees $13, 15, 15$).
- Obtained in Sage by clearing denominators in the rational win-probabilities.

One boundary curve (don't read — just to show it's not linear)

$$P_{001,000}(u,v) = u^{12}v^3 - 2u^{11}v^3 + u^{10}v^3 + u^8v^2 - 2u^7v^2 - u^6v^2 + 2u^5v^2 - u^5v - u^4v + u^2v + uv - u + v.$$

Crossing the curve $P_{001,000}(u,v) = 0$ flips the edge 001 vs 000.

What computer algebra is doing under the hood

For fixed patterns A, B, the race advantage $g_{A,B}(p)$ is a **rational function** of the bias p (finite absorbing Markov chain).

A crossover $g_{A,B}(p) = 0$ is therefore an **algebraic number**: a root of an *integer* polynomial obtained by clearing denominators.

- We certify tournaments by **exact arithmetic**: isolate roots in $(0, 1)$ and evaluate signs of the defining polynomials.
- Even in the "small" regime $\max(|A|, |B|, |C|) \leq 8$ (where we classify all cycle windows), minimal polynomials for crossover points can already have degree $>$ **100**, with very large integer coefficients.

Takeaway: the automata are small, but eliminating p leads to unexpectedly high-degree, high-height algebra.

Absurdly large polynomials for exact cross-over points

Two length-n patterns $\Rightarrow$ product automaton has n^2 transient states (plus absorbing outcomes). In an independent race, each step uses *two* tosses, so transition probabilities involve p^2, $p(1-p)$, $(1-p)^2$ (i.e. are **quadratic** in p).

- Therefore each win probability (and $g_{A,B}(p)$) is a **rational function**

$$g_{A,B}(p) = \frac{N_{A,B}(p)}{D_{A,B}(p)} \qquad \text{with} \qquad \deg N_{A,B},\ \deg D_{A,B} \ \le\ 2n^2.$$

- A cross-over $g_{A,B}(p) = 0$ reduces (after clearing denominators) to an **integer polynomial** of degree $\le 2n^2$.
- Empirically (linear-algebra computations up to $n = 30$), the *typical* **minimal polynomial degree** of a cross-over point is

$$\deg \approx 2n^2 - O(1).$$

- Example from the database: for two length-12 patterns, one cross-over had minimal degree

$$279 = 2 \cdot 12^2 - 9.$$

For $n = 30$ this heuristic suggests degrees on the order of $\sim 2 \cdot 30^2 \approx 1800$.

Takeaways

- **Same frequency, different first-hit times:** recurrence gap is $1/\mathbb{P}(T)$, but $\mathbb{E}[\tau_T]$ depends on overlap (borders).

Takeaways

- **Same frequency, different first-hit times:** recurrence gap is $1/\mathbb{P}(T)$, but $\mathbb{E}[\tau_T]$ depends on overlap (borders).
- **Fairness is exceptional:** under a fair die, stochastic dominance totally orders all patterns (and agrees with mean order).

Takeaways

- **Same frequency, different first-hit times:** recurrence gap is $1/\mathbb{P}(T)$, but $\mathbb{E}[\tau_T]$ depends on overlap (borders).
- **Fairness is exceptional:** under a fair die, stochastic dominance totally orders all patterns (and agrees with mean order).
- **Bias breaks the order:** biased coins yield incomparable waiting-time laws under stochastic dominance (simple checkpoint + mean witness).

Takeaways

- **Same frequency, different first-hit times:** recurrence gap is $1/\mathbb{P}(T)$, but $\mathbb{E}[\tau_T]$ depends on overlap (borders).
- **Fairness is exceptional:** under a fair die, stochastic dominance totally orders all patterns (and agrees with mean order).
- **Bias breaks the order:** biased coins yield incomparable waiting-time laws under stochastic dominance (simple checkpoint + mean witness).
- **Independent races add new paradoxes:** win odds are rational functions of bias; bias enables reversals and genuine non-transitive cycles.

Open problems

1. **Bias $\Leftrightarrow$ reversals?** Conjecture (coin, random tie-break): $p \neq \frac{1}{2}$ iff there exist patterns U, V with $\mathbb{E}[\tau_U] > \mathbb{E}[\tau_V]$ but $\Pr(U > V) > \frac{1}{2}$.
2. **Bias $\Rightarrow$ cycles?** Does every non-uniform i.i.d. die admit a paradoxical triple under independent races and random tie-break? *(Coin case)* For fixed $p \neq \frac{1}{2}$, what is the smallest $L_{\max}(p)$ that supports a common-bias 3-cycle?
3. **Winner flips / non-monotonicity.** For fixed patterns A, B, can $g_{A,B}(p)$ have *two or more* distinct zeros in $(0, 1)$ (i.e. can the winner flip more than once as p varies)?

Thanks!

arXiv:2601.16580

Asymptotic coverage strategy

The witnesses: For integers $n \geq 4$ and $d \in \{2, 3\}$, consider the string families

$$B_n := \mathrm{H}^n \qquad \text{vs} \qquad A_{n,d} := \mathrm{H}^{\,n-d}\,\mathrm{T}^{\,d+1}.$$

The "safe" reversal window: Let $t = 2^{-n}$ and write $p = \frac{1}{2} - \varepsilon$. Asymptotic expansion of the advantage function $g_{n,d}(p) := W^{\mathrm{rtb}}(A_{n,d} > B_n; p) - \frac{1}{2}$ yields a conservative interval $(L_{\mathrm{safe}}(n,d), R_{\mathrm{safe}}(n,d))$ where $A_{n,d}$ *wins even though it has larger mean waiting time.*

Conservative interval formulas

$$R_{\mathrm{safe}}(n,d) = \frac{1}{2} - \frac{1}{d\,2^{n+2}} \qquad \text{(mean crossover lower bound)}$$

$$L_{\mathrm{safe}}(n,d) = \frac{1}{2} - \frac{1}{d\,2^{n+1}} + \frac{n}{2d^2\,4^n} \qquad \text{(win crossover upper bound)}$$

The $+O(4^{-n})$ correction pushes the boundary inward; $(L_{\mathrm{safe}}, R_{\mathrm{safe}})$ is a rigorous subset of the true reversal window.

The "ubiquity of paradox" conjecture

Overlapping windows: The safe intervals satisfy a chaining property: the "top" of the window for $n+1$ overlaps the "bottom" of the window for n.

Coverage result (analytical + computational evidence)

Let

$$\mathcal{P}_{\mathrm{tail}} := \bigcup_{n \geq 4,\ d \in \{2,3\}} (L_{\mathrm{safe}}(n,d), R_{\mathrm{safe}}(n,d)).$$

Evidence suggests

$$\mathcal{P}_{\mathrm{tail}} \supseteq \left(\frac{1}{2} - \frac{1}{64}, \frac{1}{2}\right).$$

Consequences:

- By symmetry ($p \mapsto 1-p$, $\mathrm{H} \leftrightarrow \mathrm{T}$), the mirror interval $(\frac{1}{2}, \frac{1}{2} + \frac{1}{64})$ is covered by complemented patterns.
- Finite computation suggests the remaining "outer" regions are covered by shorter witnesses (e.g. $L \leq 8$).

Experimental claim: reversal paradoxes occur iff $p \in (0,1) \setminus \{\frac{1}{2}\}$.