361 (Tokieda's paradox)

(a) Alice tosses a coin until she sees a head followed by a tail. How many times does she toss a coin on average?

(b) Bob tosses a coin until he sees two heads in a row. How many times does he toss a coin on average?

(c) Since the probability of a head is equal to the probability of a tail, the probability of a head followed by a tail is equal to the probability of two heads in a row; each is $1/4$. So why do the answers to (a) and (b) differ?

After trying the question, scroll down to the solution.

(a) Alice tosses a coin until she sees a head followed by a tail. How many times does she toss a coin on average?

§ Before presenting the formal analysis, here are 3 helpful lemmas.

Lemma 0: $(\Sigma n{:}\ nat{+}1\cdot\ 2^{-n}) = 1$
Proof:

$$
\begin{array}{lll}
 & \Sigma n{:}\ nat{+}1\cdot\ 2^{-n} & \text{separate first term} \\
= & 1/2\ +\ \Sigma n{:}\ nat{+}2\cdot\ 2^{-n} & \text{change of variable} \\
= & 1/2\ +\ \Sigma n{:}\ nat{+}1\cdot\ 2^{-(n+1)} & \text{factor out } 1/2 \\
= & 1/2\ +\ 1/2 \times \Sigma n{:}\ nat{+}1\cdot\ 2^{-n} &
\end{array}
$$

Let $x\ =\ \Sigma n{:}\ nat{+}1\cdot\ 2^{-n}$. Then $x\ =\ 1/2\ +\ 1/2 \times x$. So $x{=}1$.

Lemma 1: $(\Sigma n{:}\ nat{+}1\cdot\ n{\times}2^{-n}) = 2$
Proof:

$$
\begin{array}{lll}
\Sigma n{:}\ nat{+}1\cdot\ n{\times}2^{-n} & & \text{separate first term} \\
= & 1/2\ +\ \Sigma n{:}\ nat{+}2\cdot\ n{\times}2^{-n} & \text{change of variable} \\
= & 1/2\ +\ \Sigma n{:}\ nat{+}1\cdot\ (n{+}1){\times}2^{-(n+1)} & \text{factor out } 1/2 \\
= & 1/2\ +\ 1/2 \times \Sigma n{:}\ nat{+}1\cdot\ (n{+}1){\times}2^{-n} & \text{separate sum} \\
= & 1/2\ +\ 1/2 \times ((\Sigma n{:}\ nat{+}1\cdot\ n{\times}2^{-n}) + (\Sigma n{:}\ nat{+}1\cdot\ 2^{-n})) & \text{Lemma 0} \\
= & 1/2\ +\ 1/2 \times ((\Sigma n{:}\ nat{+}1\cdot\ n{\times}2^{-n}) + 1) &
\end{array}
$$

Let $x\ =\ \Sigma n{:}\ nat{+}1\cdot\ n{\times}2^{-n}$. Then $x\ =\ 1/2\ +\ 1/2 \times (x{+}1)$. So $x{=}2$.

Lemma 2: $(\Sigma n{:}\ nat{+}1\cdot\ n^2 \times 2^{-n}) = 6$
Proof:

$$
\begin{array}{lll}
 & \Sigma n{:}\ nat{+}1\cdot\ n^2 \times 2^{-n} & \text{separate first term} \\
= & 1/2 + \Sigma n{:}\ nat{+}2\cdot\ n^2 \times 2^{-n} & \text{change of variable} \\
= & 1/2 + \Sigma n{:}\ nat{+}1\cdot\ (n{+}1)^2 \times 2^{-(n+1)} & \text{square, and factor out } 1/2 \\
= & 1/2 + 1/2 \times \Sigma n{:}\ nat{+}1\cdot\ (n^2 + 2{\times}n + 1) \times 2^{-n} & \text{separate sum} \\
= & 1/2 + 1/2 \times ((\Sigma n{:}\ nat{+}1\cdot\ n^2{\times}2^{-n}){+}(\Sigma n{:}\ nat{+}1\cdot\ 2{\times}n{\times}2^{-n}){+}(\Sigma n{:}\ nat{+}1\cdot\ 2^{-n})) & \\
 & & \text{Lemmas 0 and 1} \\
= & 1/2 + 1/2 \times ((\Sigma n{:}\ nat{+}1\cdot\ n^2{\times}2^{-n}) + 2{\times}2 + 1) &
\end{array}
$$

Let $x = \Sigma n{:}\ nat{+}1\cdot\ n^2{\times}2^{-n}$. Then $x\ =\ 1/2 + 1/2 \times (x + 2{\times}2 + 1)$. So $x{=}6$.

Using binary variables x and y for the results of the coin tosses, and positive natural variable n to count the tosses, here is Alice's program.

$$A\ =\ \textbf{if } 1/2 \textbf{ then } x{:=}\ head \textbf{ else } x{:=}\ tail \textbf{ fi}.\ \ n{:=}\ 1.\ \ P$$

$$
\begin{array}{ll}
P\ = & \textbf{if } 1/2 \textbf{ then } y{:=}\ head \textbf{ else } y{:=}\ tail \textbf{ fi}.\ \ n{:=}\ n{+}1. \\
 & \textbf{if } x{=}head \wedge y{=}tail \textbf{ then } ok \textbf{ else } x{:=}\ y.\ \ P \textbf{ fi}
\end{array}
$$

In the end, $x'{=}head \wedge y'{=}tail$, but we want the distribution of n' ; henceforth we ignore x' and y' . Any execution of this program produces a sequence of coin results that looks like this: $tail^*\ head^+\ tail$ where * means 0 or more, and $^+$ means 1 or more. There are $n'{-}1$ such sequences of length n' , each occurring with probability $2^{-n'}$. So I propose

$$A\ =\ (n'{-}1){\times}2^{-n'}$$

For example, $n'{=}4$ with probability $(4{-}1){\times}2^{-4}$ which is $3/16$. I need to check that A is a distribution of n' ; each value is a probability, so I must prove

$$
\begin{array}{lll}
 & (\Sigma n'\colon nat{+}1\cdot\ (n'{-}1)\times 2^{-n'}) & \text{separate sum} \\
= & (\Sigma n'\colon nat{+}1\cdot\ n'\times 2^{-n'}) - (\Sigma n'\colon nat{+}1\cdot\ 2^{-n'}) & \text{Lemmas 0 and 1} \\
= & 2-1 & \\
= & 1 &
\end{array}
$$

Now we need a hypothesis for P. If $x{=}head$ then an execution of P produces a sequence of coin tosses that looks like this: $head^*\ tail$, and there is only 1 such sequence of length $n'{-}n$, with probability $2^{-(n'-n)}$. If $x{=}tail$ then an execution of P produces a sequence of coin tosses that looks the same as an execution of A produces. So I propose

$$P \;=\; \textbf{if } x{=}head \textbf{ then } (n'{-}n \ge 1)\times 2^{-(n'-n)} \textbf{ else } (n'{-}n \ge 1)\times(n'{-}n{-}1)\times 2^{-(n'-n)} \textbf{ fi}$$

The **then**-part sums to 1 by Lemma 0, and the **else**-part is the same distribution as A. We can simplify P by factoring:

$$P \;=\; (n'{-}n \ge 1)\times 2^{-(n'-n)}\times \textbf{if } x{=}head \textbf{ then } 1 \textbf{ else } n'{-}n{-}1 \textbf{ fi}$$

Here is the proof of the A equation, starting with the right side.

$$
\begin{array}{lll}
 & \textbf{if } 1/2 \textbf{ then } x{:=}\ head \textbf{ else } x{:=}\ tail \textbf{ fi}.\ \ n{:=}\ 1.\ \ P & \text{distribute} \\
= & \textbf{if } 1/2 \textbf{ then } x{:=}\ head.\ \ n{:=}\ 1.\ \ P \textbf{ else } x{:=}\ tail.\ \ n{:=}\ 1.\ \ P \textbf{ fi} & \text{replace } P \text{ and substitutions} \\
= & \textbf{if } 1/2 \textbf{ then } (n'{-}1 \ge 1)\times 2^{-(n'-1)}\times \textbf{if } head{=}head \textbf{ then } 1 \textbf{ else } n'{-}1{-}1 \textbf{ fi} & \text{case base} \\
 & \textbf{else } (n'{-}1 \ge 1)\times 2^{-(n'-1)}\times \textbf{if } tail{=}head \textbf{ then } 1 \textbf{ else } n'{-}1{-}1 \textbf{ fi fi} & \text{case base} \\
= & \textbf{if } 1/2 \textbf{ then } (n'{\ge}2)\times 2^{-(n'-1)}\times 1 \textbf{ else } (n'{\ge}2)\times 2^{-(n'-1)}\times (n'{-}2) \textbf{ fi} & \text{factor} \\
= & (n'{\ge}2)\times 2^{-(n'-1)}\times \textbf{if } 1/2 \textbf{ then } 1 \textbf{ else } n'{-}2 \textbf{ fi} & \text{arithmetize } \textbf{if} \\
= & (n'{\ge}2)\times 2^{-(n'-1)}\times (1/2 + (n'{-}2)/2) & \text{arithmetic} \\
= & (n'{\ge}2)\times 2^{-n'}\times (n'{-}1) & \text{when } n'{=}1 \text{ result is } 0 \\
= & (n'{\ge}1)\times(n'{-}1)\times 2^{-n'} & \\
= & A &
\end{array}
$$

And now the P equation, starting with its right side.

$$
\begin{array}{lll}
 & \textbf{if } 1/2 \textbf{ then } y{:=}\ head \textbf{ else } y{:=}\ tail \textbf{ fi}.\ \ n{:=}\ n{+}1. & \\
 & \textbf{if } x{=}head \wedge y{=}tail \textbf{ then } ok \textbf{ else } x{:=}\ y.\ \ P \textbf{ fi} & \text{replace } ok \text{ and } P \\
= & \textbf{if } 1/2 \textbf{ then } y{:=}\ head \textbf{ else } y{:=}\ tail \textbf{ fi}.\ \ n{:=}\ n{+}1. & \\
 & \textbf{if } x{=}head \wedge y{=}tail \textbf{ then } n'{=}n & \\
 & \textbf{else } x{:=}\ y.\ \ (n'{-}n \ge 1)\times 2^{-(n'-n)}\times \textbf{if } x{=}head \textbf{ then } 1 \textbf{ else } n'{-}n{-}1 \textbf{ fi fi} & \text{substitution law} \\
= & \textbf{if } 1/2 \textbf{ then } y{:=}\ head \textbf{ else } y{:=}\ tail \textbf{ fi}.\ \ n{:=}\ n{+}1. & \\
 & \textbf{if } x{=}head \wedge y{=}tail \textbf{ then } n'{=}n & \\
 & \textbf{else } (n'{-}n \ge 1)\times 2^{-(n'-n)}\times \textbf{if } y{=}head \textbf{ then } 1 \textbf{ else } n'{-}n{-}1 \textbf{ fi fi} & \text{distribute} \\
= & \textbf{if } 1/2 \textbf{ then } y{:=}\ head.\ \ n{:=}\ n{+}1. & \text{substitution law twice} \\
 & \qquad\qquad \textbf{if } x{=}head \wedge y{=}tail \textbf{ then } n'{=}n & \\
 & \qquad\qquad \textbf{else } (n'{-}n \ge 1)\times 2^{-(n'-n)}\times \textbf{if } y{=}head \textbf{ then } 1 \textbf{ else } n'{-}n{-}1 \textbf{ fi fi} & \\
 & \textbf{else } y{:=}\ tail.\ \ n{:=}\ n{+}1. & \text{and twice more} \\
 & \qquad \textbf{if } x{=}head \wedge y{=}tail \textbf{ then } n'{=}n & \\
 & \qquad \textbf{else } (n'{-}n \ge 1)\times 2^{-(n'-n)}\times \textbf{if } y{=}head \textbf{ then } 1 \textbf{ else } n'{-}n{-}1 \textbf{ fi fi fi} & \\
= & \textbf{if } 1/2 \textbf{ then if } x{=}head \wedge head{=}tail \textbf{ then } n'{=}n{+}1 & \\
 & \qquad\qquad \textbf{else } (n'{-}n{-}1 \ge 1)\times 2^{-(n'-n-1)}\times \textbf{if } head{=}head \textbf{ then } 1 \textbf{ else } n'{-}n{-}1{-}1 \textbf{ fi fi} & \\
 & \textbf{else if } x{=}head \wedge tail{=}tail \textbf{ then } n'{=}n{+}1 & \\
 & \qquad \textbf{else } (n'{-}n{-}1 \ge 1)\times 2^{-(n'-n-1)}\times \textbf{if } tail{=}head \textbf{ then } 1 \textbf{ else } n'{-}n{-}1{-}1 \textbf{ fi fi fi} & \text{simplify} \\
= & \textbf{if } 1/2 \textbf{ then } (n'{-}n \ge 2)\times 2^{-(n'-n-1)} & \\
 & \textbf{else if } x{=}head \textbf{ then } n'{=}n{+}1 \textbf{ else } (n'{-}n \ge 2)\times 2^{-(n'-n-1)}\times (n'{-}n{-}2) \textbf{ fi fi} & \text{replace } \textbf{if } 1/2 \\
= & \quad 1/2\times (n'{-}n \ge 2)\times 2^{-(n'-n-1)} & \\
 & +\ 1/2\times \textbf{if } x{=}head \textbf{ then } n'{=}n{+}1 \textbf{ else } (n'{-}n \ge 2)\times 2^{-(n'-n-1)}\times (n'{-}n{-}2) \textbf{ fi} & \text{distribute} \\
= & \quad 1/2\times (n'{-}n \ge 2)\times 2^{-(n'-n-1)} & \\
 & +\ \textbf{if } x{=}head \textbf{ then } (n'{=}n{+}1)/2 \textbf{ else } (n'{-}n \ge 2)\times 2^{-(n'-n-1)}\times (n'{-}n{-}2)\ /\ 2 \textbf{ fi} & \text{arithmetic}
\end{array}
$$

$=\quad (n'-n \ge 2) \times 2^{-(n'-n)}$
$\quad + \textbf{if } x{=}head \textbf{ then } (n'{=}n{+}1)/2 \textbf{ else } (n'-n \ge 2) \times 2^{-(n'-n)} \times (n'-n-2) \textbf{ fi}$ context
$=\quad (n'-n \ge 2) \times 2^{-(n'-n)}$
$\quad + \textbf{if } x{=}head \textbf{ then } (n'{=}n{+}1) \times 2^{-(n'-n)} \textbf{ else } (n'-n \ge 2) \times 2^{-(n'-n)} \times (n'-n-2) \textbf{ fi}$ factor
$=\quad 2^{-(n'-n)} \times ((n'-n \ge 2) + \textbf{if } x{=}head \textbf{ then } n'-n = 1 \textbf{ else } (n'-n \ge 2) \times (n'-n-2) \textbf{ fi})$
check 3 cases: $n'-n \le 0$, $n'-n = 1$, and $n'-n \ge 2$
$=\quad (n'-n \ge 1) \times 2^{-(n'-n)} \times \textbf{if } x{=}head \textbf{ then } 1 \textbf{ else } n'-n-1 \textbf{ fi}$
$=\quad P$

The average value of n' is
$\quad A.\ n$
$=\quad (n'-1)\times 2^{-n'}.\ n$
$=\quad \Sigma n'': nat{+}1\cdot\ (n''-1)\times 2^{-n''} \times n''$ drop the primes (rename) and symmetry of $\times$
$=\quad \Sigma n: nat{+}1\cdot\ (n-1) \times n \times 2^{-n}$ multiply and separate sum
$=\quad (\Sigma n: nat{+}1\cdot\ n^2 \times 2^{-n}) - (\Sigma n: nat{+}1\cdot\ n \times 2^{-n})$ Lemmas 1 and 2
$=\quad 6-2$
$=\quad 4$

(b) Bob tosses a coin until he sees two heads in a row. How many times does he toss a coin on average?

§ Here is Bob's program.

$B \quad = \quad \textbf{if } 1/2 \textbf{ then } x{:=}\ head \textbf{ else } x{:=}\ tail \textbf{ fi}.\ \ n{:=}\ 1.\ \ Q$

$Q \quad = \quad \textbf{if } 1/2 \textbf{ then } y{:=}\ head \textbf{ else } y{:=}\ tail \textbf{ fi}.\ \ n{:=}\ n{+}1.$
$\quad\quad \textbf{if } x{=}head \wedge y{=}head \textbf{ then } ok \textbf{ else } x{:=}\ y.\ \ Q \textbf{ fi}$

In the end, $x'{=}head \wedge y'{=}head$, but we want the distribution of n' . Any execution of this program produces a sequence of coin results that looks like this:
$\quad *tail;\ *(head;\ tail;\ *tail);\ head;\ head$
Just to simplify the formula, if we take the final *head* and change it into an initial *tail* , we don't change the length or the probability of any sequence, and the sequences are
$\quad tail;\ *tail;\ head;\ *(tail;\ *tail;\ head)$
For each $n'{\ge}2$ there are $f\ n'$ such sequences of length n' , each occurring with probability $2^{-n'}$, where
$\quad f\ n' \ = \ \Sigma i: nat\cdot\ \ (n'-3{\times}i-1 \ge 0) \times (n'-2{\times}i-1)!\ /\ (n'-3{\times}i-1)!\ /\ i!$
The initial factor $(n'-3{\times}i-1 \ge 0)$ is equivalent to $(i < n'/3)$.
Checking: $f\,2 = 1$, $f\,3 = 1$, $f\,4 = 2$, $f\,5 = 3$.

So I propose
$\quad B \quad = \quad (n'{\ge}2) \times f\ n' \times 2^{-n'}$
For example, $n'{=}4$ with probability $(4{\ge}2) \times f\ n' \times 2^{-4}$ which is $1/8$. We need to prove that B is a distribution of n' . Each value is a probability, so we prove
$\quad \Sigma n': nat{+}1\cdot\ (n'{\ge}2) \times f\ n' \times 2^{-n'}$ drop the primes (rename)
$=\quad \Sigma n: nat{+}2\cdot\ \ f\ n \times 2^{-n}$
$=\quad \Sigma n: nat{+}2\cdot\ \ 2^{-n} \times \Sigma i: nat\cdot\ \ (n-3{\times}i-1 \ge 0) \times (n-2{\times}i-1)!\ /\ (n-3{\times}i-1)!\ /\ i!$
$=\quad$ UNFINISHED
$=\quad 1$

Now we need a hypothesis for Q , and I propose
$\quad Q \quad = \quad$ UNFINISHED

Here is the proof, starting with the B equation, right side.

$$\begin{array}{ll}
 & \textbf{if } 1/2 \textbf{ then } x{:=}\ head \textbf{ else } x{:=}\ tail \textbf{ fi}.\ \ n{:=}\ 1.\ \ Q \\
= & \text{UNFINISHED} \\
= & B
\end{array}$$

And now the Q equation, starting with its right side.

$$\begin{array}{ll}
 & \textbf{if } 1/2 \textbf{ then } y{:=}\ head \textbf{ else } y{:=}\ tail \textbf{ fi}.\ \ n{:=}\ n+1. \\
 & \textbf{if } x{=}head \wedge y{=}head \textbf{ then } ok \textbf{ else } x{:=}\ y.\ \ Q \textbf{ fi} \\
= & \text{UNFINISHED} \\
= & Q
\end{array}$$

The average value of n' is

$$\begin{array}{lll}
 & B.\ n & \text{replace } B \\
= & (n'{\ge}2) \times f\,n' \times 2^{-n'}.\ \ n & \text{definition of } . \\
= & \Sigma n'': nat{+}1\cdot\ \ (n''{\ge}2) \times f\,n'' \times 2^{-n''} \times n'' & \text{drop the primes (rename)} \\
= & \Sigma n: nat{+}2\cdot\ \ f\,n \times 2^{-n} \times n & \\
= & \Sigma n: nat{+}2\cdot\ \ n \times 2^{-n} \times \Sigma i: nat\cdot\ \ (n - 3{\times}i - 1 \ge 0) \times (n - 2{\times}i - 1)!\ /\ (n - 3{\times}i - 1)!\ /\ i! & \\
= & \text{UNFINISHED} & \\
= & 6 &
\end{array}$$

(c) Since the probability of a head is equal to the probability of a tail, the probability of a head followed by a tail is equal to the probability of two heads in a row; each is $1/4$. So why do the answers to (a) and (b) differ?

§ Alice's program in part (a) and Bob's program in part (b) are so similar that you might expect the same answer, but the answer differs. That's why Tokieda called it a paradox.