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

Lemma 0: \( (\sum n \cdot \text{nat}+1 \cdot 2^{-n}) = 1 \)
Proof:
\[
\sum n \cdot \text{nat}+1 \cdot 2^{-n} = 1/2 + \sum n \cdot \text{nat}+2 \cdot 2^{-n}
\]
\[
= 1/2 + \sum n \cdot \text{nat}+1 \cdot 2^{-(n+1)}
\]
\[
= 1/2 + 1/2 \times \sum n \cdot \text{nat}+1 \cdot 2^{-n}
\]
\[
\text{Lemma 0}
\]
Let \( x = \sum n \cdot \text{nat}+1 \cdot 2^{-n} \). Then \( x = 1/2 + 1/2 \times x \). So \( x=1 \).

Lemma 1: \( (\sum n \cdot \text{nat}+1 \cdot n \times 2^{-n}) = 2 \)
Proof:
\[
\sum n \cdot \text{nat}+1 \cdot n \times 2^{-n} = 1/2 + \sum n \cdot \text{nat}+2 \cdot n \times 2^{-n}
\]
\[
= 1/2 + \sum n \cdot \text{nat}+1 \cdot (n+1) \times 2^{-(n+1)}
\]
\[
= 1/2 + 1/2 \times \sum n \cdot \text{nat}+1 \cdot (n+1) \times 2^{-n}
\]
\[
= 1/2 + 1/2 \times ((\sum n \cdot \text{nat}+1 \cdot n \times 2^{-n}) + (\sum n \cdot \text{nat}+1 \cdot 2^{-n}))
\]
\[
\text{Lemma 0}
\]
Let \( x = \sum n \cdot \text{nat}+1 \cdot n \times 2^{-n} \). Then \( x = 1/2 + 1/2 \times (x+1) \). So \( x=2 \).

Lemma 2: \( (\sum n \cdot \text{nat}+1 \cdot n^2 \times 2^{-n}) = 6 \)
Proof:
\[
\sum n \cdot \text{nat}+1 \cdot n^2 \times 2^{-n} = 1/2 + \sum n \cdot \text{nat}+2 \cdot n^2 \times 2^{-n}
\]
\[
= 1/2 + \sum n \cdot \text{nat}+1 \cdot (n+1)^2 \times 2^{-(n+1)}
\]
\[
= 1/2 + 1/2 \times \sum n \cdot \text{nat}+1 \cdot (n^2 + 2 \times n + 1) \times 2^{-n}
\]
\[
= 1/2 + 1/2 \times ((\sum n \cdot \text{nat}+1 \cdot n^2 \times 2^{-n}) + (\sum n \cdot \text{nat}+1 \cdot 2^n \times 2^{-n} ) + (\sum n \cdot \text{nat}+1 \cdot 2^n \times 2^{-n} ))
\]
\[
\text{Lemmas 0 and 1}
\]
Let \( x = \sum n \cdot \text{nat}+1 \cdot n^2 \times 2^{-n} \). Then \( x = 1/2 + 1/2 \times (x + 2 \times x + 1) \). 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.

\[
A = \text{ if } 1/2 \text{ then } x := \text{head} \text{ else } x := \text{tail} \text{. } n := 1. \ P
\]

\[
P = \text{ if } 1/2 \text{ then } y := \text{head} \text{ else } y := \text{tail} \text{. } n := n+1.
\]
\[
\text{ if } x = \text{head} \land y = \text{tail} \text{ then } \text{ok} \text{ else } x := y. \ P \text{ fi}
\]

In the end, \( x' = \text{head} \land y' = \text{tail} \), 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: \( \text{tail} \text{ head}^* \text{ tail} \) 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
\[(\sum' \cdot n' \cdot (n' - 1) \times 2^{-n'})\]

separate sum

\[= (\sum' \cdot n' \cdot 2^{-n'}) - (\sum' \cdot n' \cdot 2^{-n'})\]

Lemmas 0 and 1

\[= 2 - 1\]

\[= 1\]

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

\[
P = \begin{cases} \text{if } x = \text{head} \text{ then } (n' - n - 1) \times 2^{-(n' - n)} \text{ else } (n' - n - 1) \times (n' - n - 1) \times 2^{-(n' - n)} & \text{fi} \end{cases} \]

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:

\[
P = (n' - n - 1) \times 2^{-(n' - n)} \times \text{if } x = \text{head} \text{ then } 1 \text{ else } n' - n - 1 \text{ fi}
\]

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

if 1/2 then \( x := \text{head} \text{ else } x := \text{tail} \text{ fi. } n := 1 \text{. } P\) distribute

= if 1/2 then \( x := \text{head} \text{. } n := 1 \). \( P \text{ else } x := \text{tail} \text{. } n := 1 \). \( P \text{ fi replace } P \text{ and substitutions

= if 1/2 then \((n' - n - 1) \times 2^{-(n' - n)} \times \text{if } head = \text{head} \text{ then } 1 \text{ else } n' - 1 - 1 \text{ fi}

\text{case base}\)

else \((n' - 1 \times 2^{-(n' - 1)} \times \text{if } tail = \text{head} \text{ then } 1 \text{ else } n' - 1 - 1 \text{ fifi case base

= if 1/2 then \((n' \geq 2) \times 2^{-(n' - 1)} \times 1 \text{ else } (n' \geq 2) \times 2^{-(n' - 1)} \times (n' - 2) \text{ fi}

\text{factor}\)

= \((n' \geq 2) \times 2^{-(n' - 1)} \times \text{if } 1/2 \text{ then } 1 \text{ else } n' - 2 \text{ fi

\text{arithmetize if}\)

= \((n' \geq 2) \times 2^{-(n' - 1)} \times (1/2 + (n' - 2)/2)

\text{arithmetic}\)

= \((n' \geq 2) \times 2 - n \times (n' - 1) \text{ when } n' = 1 \text{ result is 0}

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

= A

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

if 1/2 then \( y := \text{head} \text{ else } y := \text{tail} \text{ fi. } n := n + 1 \).

if \( x = \text{head} \land y = \text{tail then ok else x := y. } P \text{ fi replace ok and } P

= if 1/2 then \( y := \text{head} \text{ else } y := \text{tail} \text{. } n := n + 1 \).

if \( x = \text{head} \land y = \text{tail then n' = n

\text{substitution law}\)

else \( x := y. \ (n' - n + 1) \times 2^{-(n' - n)} \times \text{if } x = \text{head then } 1 \text{ else } n' - n - 1 \text{ fi fi substitution law

= if 1/2 then \( y := \text{head} \text{ else } y := \text{tail} \text{. } n := n + 1 \).

if \( x = \text{head} \land y = \text{tail then n' = n

\text{substitution law twice}\)

else \( (n' - n + 1) \times 2^{-(n' - n)} \times \text{if } y = \text{head then } 1 \text{ else } n' - n - 1 \text{ fi fi distribute

= if 1/2 then \( y := \text{head} \text{. } n := n + 1 \).

\text{substitution law twice and twice more}\)

if \( x = \text{head} \land y = \text{tail then n' = n

\text{else } \text{substitution law}\)

else \( (n' - n + 1) \times 2^{-(n' - n)} \times \text{if } y = \text{head then } 1 \text{ else } n' - n - 1 \text{ fi fi fi

= if 1/2 then \( x = \text{head} \land head = \text{tail then n' = n + 1

\text{else } \text{substitution law}\)

else \( (n' - n + 1) \times 2^{-(n' - n)} \times \text{if } head = \text{head then } 1 \text{ else } n' - n - 1 - 1 \text{ fi fi

else if \( x = \text{head} \land tail = \text{tail then n' = n + 1

\text{else } \text{substitution law}\)

else \( (n' - n - 1) \times 2^{-(n' - n)} \times \text{if } tail = \text{head then } 1 \text{ else } n' - n - 1 - 1 \text{ fi fi fi

simplify

= if 1/2 then \( (n' - n - 1) \geq 1 \times 2^{-(n' - n)} \times \text{if } tail = \text{head then } 1 \text{ else } n' - n - 1 - 1 \text{ fi fi fi

\text{replace if } 1/2

= 1/2 \times (n' - n \geq 2) \times 2^{-(n' - n)} \times 1 \text{ else } (n' - n \geq 2) \times 2^{-(n' - n)} \times (n' - n - 2) \text{ fi fi arithmetic

+ 1/2 \times \text{if } x = \text{head then } n' = n + 1 \text{ else } (n' - n \geq 2) \times 2^{-(n' - n)} \times (n' - n - 2) \text{ fi fi distribute

= 1/2 \times (n' - n \geq 2) \times 2^{-(n' - n)} \times 1 \text{ else } (n' - n \geq 2) \times 2^{-(n' - n)} \times (n' - n - 2) / 2 \text{ fi arithmetic
Here is Bob's program.

\[ B = \text{if } x = \text{head} \text{ then } n' = n + 1 \text{ else } (n' - n) \times 2^{-(n' - n)} \times (n' - n - 2) \]

\[ Q = \text{if } x = \text{head} \text{ then } n' - n = 1 \text{ else } (n' - n) \times 2^{-(n' - n)} \]

The average value of \( n' \) is

\[ A_n = \frac{1}{n} \]

\[ \sum n''': \text{nat} \times (n'' - 1) \times 2^{-n''} \times n'' \]

\[ \sum n: \text{nat} \times (n - 1) \times n \times 2^{-n} \]

\[ \sum n: \text{nat} \times n^2 \times 2^{-n} - (\sum n: \text{nat} \times n \times 2^{-n}) \]

\[ = 6 - 2 \]

\[ = 4 \]

(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.

\[ B = \text{if } 1/2 \text{ then } x := \text{head} \text{ else } x := \text{tail} \text{. } n := 1 \text{. } Q \]

\[ Q = \text{if } 1/2 \text{ then } y := \text{head} \text{ else } y := \text{tail} \text{. } n := n + 1 \text{. } \]

\[ \text{if } x = \text{head} \land y = \text{head} \text{ then } \text{ok} \text{ else } x := y \text{. } Q \text{ fi} \]

In the end, \( x' = \text{head} \land y' = \text{head} \), but we want the distribution of \( n' \). Any execution of this program produces a sequence of coin results that looks like this:

*tail; *(head; tail; *tail); head; head

Just to simplify the formula, if we take the final \text{head} and change it into an initial \text{tail}, we don't change the length or the probability of any sequence, and the sequences are

\[ \text{tail; *tail; head; *(tail; *tail; head)} \]

For each \( n' \geq 2 \) there are \( f n' \) such sequences of length \( n' \), each occurring with probability \( 2^{-n'} \), where

\[ f n' = \sum i: \text{nat} \times (n' - 3 \times i - 1 \geq 0) \times (n' - 2 \times i - 1)! / (n' - 3 \times i - 1)! / i! \]

The initial factor \( (n' - 3 \times i - 1 \geq 0) \) is equivalent to \( i < n'/3 \).

Checking: \( f 2 = 1 \text{, } f 3 = 1 \text{, } f 4 = 2 \text{, } f 5 = 3 \).

So I propose

\[ B = (n' \geq 2) \times f n' \times 2^{-n} \]

For example, \( n' = 4 \) with probability \( (4 \geq 2) \times f 4 \times 2^{-4} \) which is \( 1/8 \). We need to prove that \( B \) is a distribution of \( n' \). Each value is a probability, so we prove

\[ \sum n': \text{nat} \times f n' \times 2^{-n} \]

\[ \sum n: \text{nat} \times f n \times 2^{-n} \]

\[ \sum n: \text{nat} \times 2^{-n} \times \sum i: \text{nat} \times (n' - 3 \times i - 1 \geq 0) \times (n - 2 \times i - 1)! / (n - 3 \times i - 1)! / i! \]

\[ = \text{UNFINISHED} \]

Now we need a hypothesis for \( Q \), and I propose

\[ Q = \text{UNFINISHED} \]
Here is the proof, starting with the $B$ equation, right side.

```plaintext
if 1/2 then $x := \text{head}$ else $x := \text{tail}$ fi. $n := 1$. $Q$
```

```plaintext
= UNFINISHED
= $B$
```

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

```plaintext
if 1/2 then $y := \text{head}$ else $y := \text{tail}$ fi. $n := n+1$.
if $x=\text{head} \land y=\text{head}$ then ok else $x := y$. $Q$ fi
```

```plaintext
= UNFINISHED
= $Q$
```

The average value of $n'$ is

```plaintext
B. n
= (n'\geq 2) \times f n' \times 2^{-n'}, n
= \sum_{n' : \text{nat}} n' + 1 \cdot (n' \geq 2) \times f n' \times 2^{-n'} \times n''
= \sum_{n : \text{nat}} n + 2 \cdot f n \times 2^{-n} \times n
= \sum_{n : \text{nat}} n \times 2^{-n} \times \sum_{i : \text{nat}} (n - 3 \times i - 1 \geq 0) \times (n - 2 \times i - 1)! / (n - 3 \times i - 1)! / i!
= UNFINISHED
= 6
```

(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.