

346 (Mr.Bean's socks) Mr.Bean is trying to get a matching pair of socks from a drawer containing an inexhaustible supply of red and blue socks. He begins by withdrawing two socks at random. If they match, he is done. Otherwise, he throws away one of them at random, withdraws another sock at random, and repeats. How long will it take him to get a matching pair? Assume that a sock withdrawn from the drawer has $1/2$ probability of being each color, and that a sock that is thrown away also has a $1/2$ probability of being each color.

§ Informally, here is Mr.Bean's program.

```

choose a sock color with the left hand.
choose a sock color with the right hand.
loop

```

where

```

loop  $\Leftarrow$  if sock colors match then ok
           else choose a hand and a sock color for that hand. loop fi

```

Let variables L and R represent the color of socks held in Mr.Bean's left and right hands, and let time variable t count iterations. Formally, the program is

```

if rand 2 then  $L := red$  else  $L := blue$  fi.
if rand 2 then  $R := red$  else  $R := blue$  fi.
loop

```

where

```

loop  $\Leftarrow$  if  $L=R$  then ok
           else if rand 2 then if rand 2 then  $L := red$  else  $L := blue$  fi
                 else if rand 2 then  $R := red$  else  $R := blue$  fi fi.
            $t := t+1$ . loop fi

```

All occurrences of *rand* 2 can be replaced by $1/2$. Since *red* and *blue* are the only two values for L and R , the two initialization lines can be simplified as follows:

$$\begin{aligned}
& \mathbf{if} \ 1/2 \ \mathbf{then} \ L := red \ \mathbf{else} \ L := blue \ \mathbf{fi}. \\
& \mathbf{if} \ 1/2 \ \mathbf{then} \ R := red \ \mathbf{else} \ R := blue \ \mathbf{fi} \qquad \text{replace } \mathbf{if} \ \text{and } := \\
= & \ 1/2 \times (L'=red) \times (R'=R) \times (t'=t) + 1/2 \times (L'=blue) \times (R'=R) \times (t'=t). \\
& \ 1/2 \times (R'=red) \times (L'=L) \times (t'=t) + 1/2 \times (R'=blue) \times (L'=L) \times (t'=t) \\
= & \ ((L'=red) + (L'=blue)) \times (R'=R) \times (t'=t) / 2. \\
& \ ((R'=red) + (R'=blue)) \times (L'=L) \times (t'=t) / 2 \\
& \qquad \text{for either value of } L' \text{ the sum is } 1, \text{ and similarly for } R' \\
= & \ (R'=R) \times (t'=t) / 2. \ (L'=L) \times (t'=t) / 2 \qquad \text{replace } . \ \text{and use one-point law} \\
= & \ (t'=t)/4
\end{aligned}$$

Similarly the loop body can be simplified:

$$\begin{aligned}
& \mathbf{if} \ 1/2 \ \mathbf{then} \ \mathbf{if} \ 1/2 \ \mathbf{then} \ L := red \ \mathbf{else} \ L := blue \ \mathbf{fi} \\
& \ \mathbf{else} \ \mathbf{if} \ 1/2 \ \mathbf{then} \ R := red \ \mathbf{else} \ R := blue \ \mathbf{fi} \ \mathbf{fi} \\
= & \ ((L'=L) + (R'=R)) \times (t'=t) / 4
\end{aligned}$$

The program is now

```

( $t'=t$ )/4. loop

```

where

```

loop  $\Leftarrow$  if  $L=R$  then ok else  $((L'=L) + (R'=R)) \times (t'=t) / 4$ .  $t := t+1$ . loop fi

```

After three failed attempts to define *loop* I propose

```

loop = if  $L=R$  then ok else  $(L'=R') \times (t'>t) \times 2^{t-t'}$  fi

```

Here's the proof of the refinement, starting with the right side.

$$\begin{aligned}
& \mathbf{if} \ L=R \ \mathbf{then} \ \mathbf{ok} \ \mathbf{else} \ ((L'=L) + (R'=R)) \times (t'=t) / 4. \ t := t+1. \ \mathbf{loop} \ \mathbf{fi} \qquad \text{expand } \mathbf{loop} \\
= & \ \mathbf{if} \ L=R \ \mathbf{then} \ \mathbf{ok} \ \mathbf{else} \ ((L'=L) + (R'=R)) \times (t'=t) / 4. \ t := t+1. \\
& \qquad \mathbf{if} \ L=R \ \mathbf{then} \ \mathbf{ok} \ \mathbf{else} \ (L'=R') \times (t'>t) \times 2^{t-t'} \ \mathbf{fi} \ \mathbf{fi} \qquad \text{expand last } \mathbf{ok} \\
= & \ \mathbf{if} \ L=R \ \mathbf{then} \ \mathbf{ok} \\
& \ \mathbf{else} \ ((L'=L) + (R'=R)) \times (t'=t) / 4. \ t := t+1.
\end{aligned}$$

$$\begin{aligned}
& \text{if } L=R \text{ then } (L'=L) \times (R'=R) \times (t'=t) \text{ else } (L'=R') \times (t'>t) \times 2^{t-t'} \text{ fi fi substitution} \\
= & \text{if } L=R \text{ then } ok \\
& \text{else } ((L'=L) + (R'=R)) \times (t'=t) / 4. \\
& \text{if } L=R \text{ then } (L'=L) \times (R'=R) \times (t'=t+1) \text{ else } (L'=R') \times (t'>t+1) \times 2^{t+1-t'} \text{ fi fi} \\
& \hspace{15em} \text{sequential composition}
\end{aligned}$$

$$\begin{aligned}
= & \text{if } L=R \text{ then } ok \\
& \text{else } \Sigma L'', R'', t'' \cdot ((L''=L) + (R''=R)) \times (t''=t) / 4 \times \\
& \text{if } L''=R'' \text{ then } (L''=L'') \times (R''=R'') \times (t''=t''+1) \text{ else } (L''=R'') \times (t''>t''+1) \times 2^{t''+1-t''} \text{ fi fi} \\
& \hspace{15em} \text{expand if}
\end{aligned}$$

$$\begin{aligned}
= & \text{if } L=R \text{ then } ok \\
& \text{else } \Sigma L'', R'', t'' \cdot ((L''=L) + (R''=R)) \times (t''=t) / 4 \times \\
& \quad ((L''=R'') \times (L''=L'') \times (R''=R'') \times (t''=t''+1) \\
& \quad + (L'' \neq R'') \times (L''=R'') \times (t''>t''+1) \times 2^{t''+1-t''}) \text{ fi} \\
& \hspace{15em} \text{distribute}
\end{aligned}$$

$$\begin{aligned}
= & \text{if } L=R \text{ then } ok \\
& \text{else } (\Sigma L'', R'', t'' \cdot (L''=L) \times (t''=t) / 4 \times (L''=R'') \times (L''=L'') \times (R''=R'') \times (t''=t''+1)) \\
& + (\Sigma L'', R'', t'' \cdot (L''=L) \times (t''=t) / 4 \times (L'' \neq R'') \times (L''=R'') \times (t''>t''+1) \times 2^{t''+1-t''}) \\
& + (\Sigma L'', R'', t'' \cdot (R''=R) \times (t''=t) / 4 \times (L''=R'') \times (L''=L'') \times (R''=R'') \times (t''=t''+1)) \\
& + (\Sigma L'', R'', t'' \cdot (R''=R) \times (t''=t) / 4 \times (L'' \neq R'') \times (L''=R'') \times (t''>t''+1) \times 2^{t''+1-t''}) \text{ fi}
\end{aligned}$$

All the L and R variables, with any number of primes, are two-valued.
In the **else** context, $L \neq R$. So in the second line, R'' can only be R .
Similarly, in the last line, L'' can only be L .
So we can one-point all double-primed variables.

$$\begin{aligned}
= & \text{if } L=R \text{ then } ok \\
& \text{else } (L'=R'=L) \times (t'=t+1) / 4 \\
& + (L'=R') \times (t'>t+1) \times 2^{t+1-t'} / 4 \\
& + (L'=R'=R) \times (t'=t+1) / 4 \\
& + (L'=R') \times (t'>t+1) \times 2^{t+1-t'} / 4 \text{ fi}
\end{aligned}$$

In the **else** context, $L \neq R$, so exactly one of the first and third lines is nonzero.
The second and fourth lines are identical, and can be added.

$$\begin{aligned}
= & \text{if } L=R \text{ then } ok \\
& \text{else } (L'=R') \times (t'=t+1) / 2 \\
& + (L'=R') \times (t'>t+1) \times 2^{t+1-t'} / 2 \text{ fi}
\end{aligned}$$

Where $t'=t+1$, dividing by 2 is the same as multiplying by $2^{t-t'}$.
And simplify the last line.

$$\begin{aligned}
= & \text{if } L=R \text{ then } ok \\
& \text{else } (L'=R') \times (t'=t+1) \times 2^{t-t'} \\
& + (L'=R') \times (t'>t+1) \times 2^{t-t'} \text{ fi} \\
& \hspace{15em} \text{now combine the two terms}
\end{aligned}$$

$$\begin{aligned}
= & \text{if } L=R \text{ then } ok \text{ else } (L'=R') \times (t'>t) \times 2^{t-t'} \text{ fi} \\
= & \text{loop}
\end{aligned}$$

Now we put the initialization together with the loop distribution to calculate the final state distribution.

$$\begin{aligned}
& (t'=t)/4. \text{ loop} \hspace{15em} \text{omitting several steps} \\
= & (L'=R') \times (t' \geq t) \times 2^{t-t'-1} \\
& \text{The average value of } t' \text{ is} \\
& (L'=R') \times (t' \geq t) \times 2^{t-t'-1}. \text{ } t \hspace{15em} \text{omitting several steps} \\
= & t+1
\end{aligned}$$

On average, Mr.Bean draws the initial two socks plus one more sock from the drawer.