BCB BCB/GDCB/STAT/COM S 568 Spring 2010

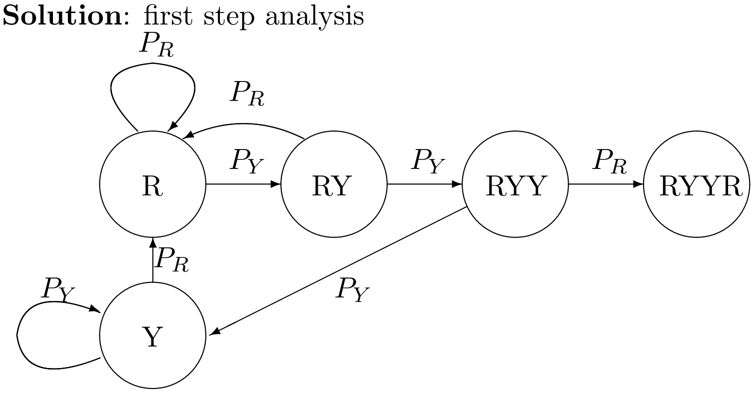
Homework 2 Solutions

January 26, 2010

Due one week later. Answers to selected problems will be posted.

1) Represent a sequence over the alphabet R, Y as a Markov chain with states R and Y, with initial and transition probabilities all equal to 0.5. What is the expected waiting time to find the word "RYYR"? Please use two independent methods (first step analysis and renewal theory application) to calculate this probability and check your result.

Solution: first step analysis



$$\begin{cases} \mu_Y &= 1 + P_Y * \mu_Y + P_R * \mu_R \\ \mu_R &= 1 + P_Y * \mu_{RY} + P_R * \mu_R \\ \mu_{RY} &= 1 + P_Y * \mu_{RYY} + P_R * \mu_R \\ \mu_{RYY} &= 1 + P_Y * \mu_Y + P_R * \mu_{RYYR} \\ \mu_{RYYR} &= 0 \end{cases} \Longrightarrow \begin{cases} \mu_Y &= 1 + \frac{1}{2} * \mu_Y + \frac{1}{2} * \mu_R \\ \mu_R &= 1 + \frac{1}{2} * \mu_{RY} + \frac{1}{2} * \mu_R \\ \mu_{RY} &= 1 + \frac{1}{2} * \mu_{RYY} + \frac{1}{2} * \mu_R \\ \mu_{RYY} &= 1 + \frac{1}{2} * \mu_Y + \frac{1}{2} * \mu_{RYYR} \\ \mu_{RYYR} &= 0 \end{cases}$$

$$\Longrightarrow \begin{cases} \mu_Y &= 18 \\ \mu_R &= 16 \\ \mu_{RY} &= 14 \\ \mu_{RYY} &= 10 \\ \mu_{RYYR} &= 0 \end{cases} \Longrightarrow \mu = 1 + \frac{1}{2} * \mu_Y + \frac{1}{2} * \mu_R = 18$$

Solution: renewal theory application

———RYYR

—RYYRYYR

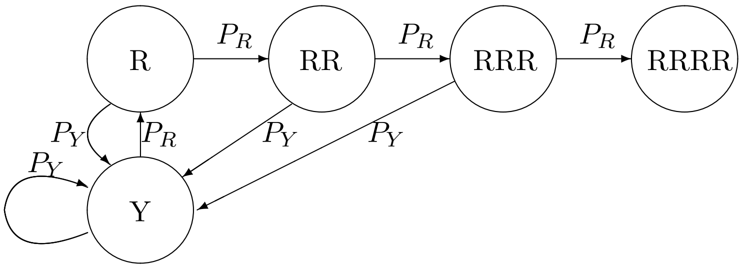
$P_R^2 * P_Y^2 = u_{n-3} * P_R * P_Y^2 + u_n$, where u_n is the probability of the pattern occurrning at position n.

When n $\rightarrow \infty$, $u_{n-3} = u_n = \frac{1}{\mu}$.

Since $P_R = P_Y = \frac{1}{2} \Longrightarrow \frac{1}{16} = \frac{1}{\mu} * \frac{1}{8} + \frac{1}{\mu} \Longrightarrow \mu = 18$

2) Solve the problems posted in (1) for the word "RRRR" instead of "RYYR".

Solution: first step analysis



$$\begin{cases} \mu_Y & = 1 + P_Y * \mu_Y + P_R * \mu_R \\ \mu_R & = 1 + P_Y * \mu_Y + P_R * \mu_{RR} \\ \mu_{RR} & = 1 + P_Y * \mu_Y + P_R * \mu_{RRR} \\ \mu_{RRR} & = 1 + P_Y * \mu_Y + P_R * \mu_{RRRR} \\ \mu_{RRRR} & = 0 \end{cases} \Longrightarrow \begin{cases} \mu_Y & = 1 + \frac{1}{2} * \mu_Y + \frac{1}{2} * \mu_R \\ \mu_R & = 1 + \frac{1}{2} * \mu_Y + \frac{1}{2} * \mu_{RR} \\ \mu_{RR} & = 1 + \frac{1}{2} * \mu_Y + \frac{1}{2} * \mu_{RRR} \\ \mu_{RRR} & = 1 + \frac{1}{2} * \mu_Y + \frac{1}{2} * \mu_{RRRR} \\ \mu_{RRRR} & = 0 \end{cases}$$

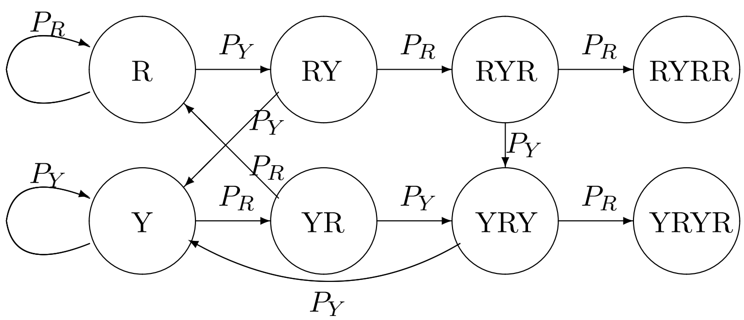
$$\Longrightarrow \begin{cases} \mu_Y & = 30 \\ \mu_R & = 28 \\ \mu_{RR} & = 24 \\ \mu_{RRR} & = 16 \\ \mu_{RRRR} & = \ 0 \end{cases} \Longrightarrow \mu = 1 + \frac{1}{2} * \mu_Y + \frac{1}{2} * \mu_R = 30$$

Solution: renewal theory application

————RRRR
———RRRRR
——RRRRRR
—RRRRRRR
$P_R^4 = u_{n-3} * P_R^3 + u_{n-2} * P_R^2 + u_{n-1} * P_R + u_n$
When n $\rightarrow \infty$, $u_{n-3} = u_{n-2} = u_{n-1} = u_n = \frac{1}{\mu}$.
Since $P_R = \frac{1}{2} \Longrightarrow \frac{1}{16} = \frac{1}{\mu} * (\frac{1}{8} + \frac{1}{4} + \frac{1}{2} + 1) \Longrightarrow \mu = 30$

3) Under the same conditions as in (1) and (2), what is the probability that the word "RYRR" occurs before "YRYR"? How about "YRYY" versus "YRYR"?

Solution: "RYRR" versus "YRYR"

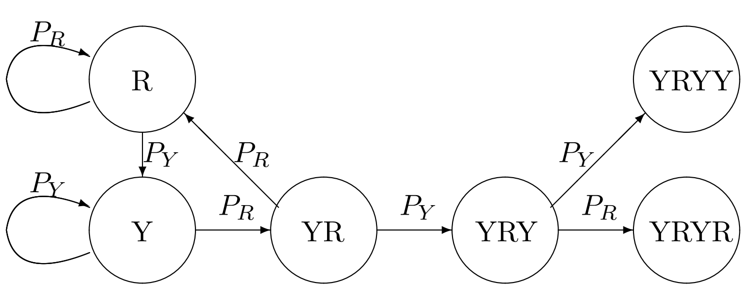


Let ρ_R be the probability that the word "RYRR" occurs before "YRYR" when the current state is "R". Then

$$\begin{cases} \rho_R &= P_R*\rho_R + P_Y*\rho_{RY} \\ \rho_{RY} &= P_R*\rho_{RYR} + P_Y*\rho_Y \\ \rho_{RYR} &= P_R*\rho_{RYRR} + P_Y*\rho_{YRY} \\ \rho_{RYRR} &= 1 \\ \rho_Y &= P_Y*\rho_Y + P_R*\rho_{YR} \\ \rho_{YR} &= P_Y*\rho_{YRY} + P_R*\rho_R \\ \rho_{YRY} &= P_Y*\rho_Y + P_R*\rho_{YRYR} \\ \rho_{YRYR} &= 0 \end{cases} \implies \begin{cases} \rho_R &= \frac{1}{2}*\rho_R + \frac{1}{2}*\rho_{RY} \\ \rho_{RY} &= \frac{1}{2}*\rho_{RYR} + \frac{1}{2}*\rho_Y \\ \rho_{RYR} &= \frac{1}{2}*\rho_{RYRR} + \frac{1}{2}*\rho_{YRY} \\ \rho_{RYRR} &= 1 \\ \rho_Y &= \frac{1}{2}*\rho_Y + \frac{1}{2}*\rho_{YR} \\ \rho_{YR} &= \frac{1}{2}*\rho_{YRY} + \frac{1}{2}*\rho_R \\ \rho_{YRY} &= \frac{1}{2}*\rho_Y + \frac{1}{2}*\rho_{YRYR} \\ \rho_{YRYR} &= 0 \end{cases}$$

$$\implies \begin{cases} \rho_Y &= \frac{2}{7} \\ \rho_R &= \frac{3}{7} \\ \rho_{YR} &= \frac{2}{7} \\ \rho_{RY} &= \frac{3}{7} \\ \rho_{RYR} &= \frac{4}{7} \\ \rho_{YRY} &= \frac{1}{7} \end{cases} \implies \rho = \frac{1}{2}*\rho_Y + \frac{1}{2}*\rho_R = \frac{1}{2}*(\frac{2}{7}+\frac{3}{7}) = \frac{5}{14}$$

Solution: "YRYY" versus "YRYR"



There are two outcomes from pattern "YRY", which are "YRYY" and "YRYR". The probability to add a "Y" or a "R" is the same. Therefore $\rho = \frac{1}{2}$.

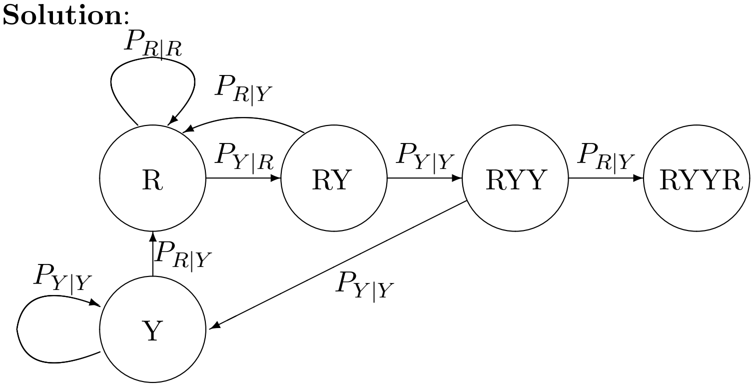
4) Represent a sequence over the alphabet R, Y as a Markov chain with states R and Y, with initial and transition probabilities as follows:

$$\textbf{The initial probability} \begin{cases} P(Y) = 0.4 \\ P(R) = 0.6 \end{cases}$$

$$\textbf{The transation probability} \begin{cases} P(Y|Y) = 0.3 \\ P(R|Y) = 0.7 \\ P(Y|R) = 0.8 \\ P(R|R) = 0.2 \end{cases}$$

What's the expected waiting time to find the word "RYYR".

Solution:



$$\begin{cases} \mu_Y &= 1 + P_{Y|Y} * \mu_Y + P_{R|Y} * \mu_R \\ \mu_R &= 1 + P_{Y|R} * \mu_{RY} + P_{R|R} * \mu_R \\ \mu_{RY} &= 1 + P_{Y|Y} * \mu_{RYY} + P_{R|Y} * \mu_R \\ \mu_{RYY} &= 1 + P_{Y|Y} * \mu_Y + P_{R|Y} * \mu_{RYYR} \\ \mu_{RYYR} &= 0 \end{cases} \implies \begin{cases} \mu_Y &= 1 + 0.3 * \mu_Y + 0.7 * \mu_R \\ \mu_R &= 1 + 0.8 * \mu_{RY} + 0.2 * \mu_R \\ \mu_{RY} &= 1 + 0.3 * \mu_{RYY} + 0.7 * \mu_R \\ \mu_{RYY} &= 1 + 0.3 * \mu_Y + 0.7 * \mu_{RYYR} \\ \mu_{RYYR} &= 0 \end{cases}$$

$$\implies \begin{cases} \mu_Y &= \frac{695}{49} = 14.18 \\ \mu_R &= \frac{625}{49} = 12.76 \\ \mu_{RY} &= \frac{2255}{196} = 11.51 \\ \mu_{RYY} &= \frac{515}{98} = 5.26 \\ \mu_{RYYR} &= 0 \end{cases} \implies$$

$$\implies \mu = 1 + P_Y * \mu_Y + P_R * \mu_R = 1 + 0.4 * \frac{695}{49} + 0.6 * \frac{625}{49} = \frac{702}{49} = 14.33$$

5) Represent a sequence over the alphabet R, Y as a Markov chain with states R and Y, with initial and transition probabilities all equal to 0.5. What is the expected waiting time to find the pattern "RYYR" twice? Calculate under two different conditions: 1) The two occurences are counted even if they overlap; 2) The two occurences are required to be non-overlapping. Compare the results with your answer to (1).

Solution: 1) The two occurences are counted even if they overlap
From the result of (1), we know that the expected waiting time to find the pattern "RYYR" once is 18.

$$\begin{cases} \mu_Y & = 18 \\ \mu_R & = 16 \\ \mu_{RY} & = 14 \\ \mu_{RYY} & = 10 \\ \mu_{RYYR} & = 0 \end{cases} \Longrightarrow \mu = 1 + \frac{1}{2} * \mu_Y + \frac{1}{2} * \mu_R = 18$$

If the two occurences are counted even if they overlap, then for the second occurence, we can use μ_R. Here "R" is the last letter in the first occurence of "RYYR", which can also be used as the first letter of the second occurence. Therefore $\mu_{twice,overlap} = \mu + \mu_R = 34$. $\mu_{twice,overlap}$ is less than 2*μ.

Solution: 2) The two occurences are required to be non-overlapping
If the two occurences are required to be non-overlapping, then $\mu_{twice,non-overlapping} = \mu + \mu = 2 * \mu$ = 36.