HS 2013 Prof. R. Wattenhofer / K.-T. Foerster, T. Langner, J. Seidel

Discrete Event Systems

Solution to Exercise Sheet 8

1 PageRank

a) With $v = \begin{pmatrix} 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 \end{pmatrix}$ we get a PageRank vector v of $\begin{pmatrix} 1 & 2 & 2 & 0 \end{pmatrix}$.

Notice that website v_4 is quite important for the PageRank, even though it is just a collection of links!

b) We first calculate $d_1 = 1, d_2 = 1, d_3 = 0, d_4 = 3$ and now get a PageRank vector of

$$v = \begin{pmatrix} 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 1/3 & 1/3 & 1/3 & 0 \end{pmatrix} = \begin{pmatrix} 1/3 & 4/3 & 4/3 & 0 \end{pmatrix}$$

We now reduced the importance of the link–collection v_4, but still, we do not account for the fact that v_4 is not important at all – nobody recommends it!

c) The results of your iterations should look like this:

(i) $\begin{pmatrix} 1/3 & 4/3 & 4/3 & 0 \end{pmatrix}$

(ii) $\begin{pmatrix} 0 & 1/3 & 4/3 & 0 \end{pmatrix}$

(iii) $\begin{pmatrix} 0 & 0 & 1/3 & 0 \end{pmatrix}$

(iv) $\begin{pmatrix} 0 & 0 & 0 & 0 \end{pmatrix}$

(v) $\begin{pmatrix} 0 & 0 & 0 & 0 \end{pmatrix}$

As you can see, without the "Random Surfer" we run into problems – the Markov Chain is not ergodic! However, continuing these calculations with a "Random Surfer" might get a bit tedious – so let us look at the other exercises now :-)

2 Colour Blindness

a) The number of colour blind people X in a sample of 100 is binomially distributed where each person is colour blind with probability $p = 0.02$. Hence we have (slide 4/15)

$$\Pr[X = k] = \binom{100}{k} p^k (1-p)^{100-k} \ .$$

The probability that at most one person out of 100 is colour blind is given by

$$\begin{aligned}\Pr[X \leq 1] &= \Pr[X = 0] + \Pr[X = 1] \\ &= \binom{100}{0} p^0 (1-p)^{100} + \binom{100}{1} p^1 (1-p)^{100-1} \\ &\approx 0.403\end{aligned}$$

If we assume X to be Poisson-distributed (see **b)**), we get (slide 4/15)

$$\Pr[X \leq 1] \approx 0.406 \ .$$

b) Since the sample size n is large and the probability for someone being colour blind is small, we can estimate the distribution of colour blind people with the Poisson distribution.

The Poisson distribution

The Poisson distribution is a *discrete* probability distribution which is applied often to approximate the binomial distribution for large number n of repetitions and small success probability p of the underlying Bernoulli experiments. According to two frequently used rules of thumb, this approximation is good if $n \geq 20$ and $p \leq 0.05$, or if $n \geq 100$ and $np \leq 10$.

The Poisson distribution is often used to estimate the probability of a given number of events occurring in a fixed interval of time and/or space if these events occur with a known average rate and independently of the time since the last event. The parameter $\lambda = np$ of the distribution is the expected number of occurrences in the interval.

$$\Pr[X = x] = \frac{\lambda^x}{x!} e^{-\lambda}$$

Since we expect the sample size n to be larger than 20 and we have $p = 0.02$, we can assume the number X of colour blind persons in a sample of n persons to be Poisson-distributed with parameter $\lambda = np = n/50$. The probability that at least one person is colour blind in a sample of size n is now given by

$$\begin{aligned}\Pr[X \geq 1] &= 1 - \Pr[X = 0] \\ &= 1 - e^{-\lambda} \cdot \frac{\lambda^0}{0!} \\ &= 1 - e^{-n/50} \ .\end{aligned}$$

Solving the inequality $\Pr[X \geq 1] \geq 90\%$ for n yields $n \geq 116$. Hence, in a sample of 116 persons we have at least one colour blind person with probability 90%.