# Problem Set 3

**Problem 1.** For the Bayesian network described in slide 11 of lecture 5, mark the following questions as true or false.

1. $P(f|d) = P(f)$.
2. $P(n|d, f) = P(n|f)$.
3. $n$ and $g$ are independent.
4. $t$ and $f$ are independent.
5. $t$ is independent of $d$, given $g$.

**Answer.**

1. True (f prior probability)
2. True (n is independent of d)
3. False (n is independent of g, given f but not otherwise)
4. False (f is an ancestor of t)
5. True (d's influence on t is via g)

**Problem 2.** Once again referring to the Bayesian shown in slide 11 of lecture 5, compute $P(f|t)$.

**Answer.**

```
  P(f|t) = P(t|f)P(f)/P(t) = 0.3P(t|f)/P(t)        (1)
Similarly,
  P(^f|t) = 0.7P(t|^f)/P(t)                         (2)

  P(t|f) = P(t|g)P(g|f) + P(t|^g)P(^g|f)
         = 0.99 P(g|f) + 0.1P(^g|f)                 (3)
Similarly,
  P(t|^f) = 0.99P(g|^f) + 0.1P(^g|^f)               (4)

   P(g|f) = P(g|fd)P(d) + P(g|f^d)P(^d)
          = 0.99*.4 + 0.8*0.6 = 0.876               (5)
Similarly,
   P(g|^f) = 0.9*0.4+0.3*0.6=0.54                   (6)
   P(^g|f) = 0.01*.4+0.2*.6=0.124                   (7)
   P(^g|^f)=0.1*0.4+0.7*0.6=.46                     (8)

(5)..(8) in (3) and (4) gives
   P(t|f) ~ 0.88  and  P(t|^f) ~ 0.58
Sub. in (1) and (2) and using
```

```
  P(f|t) + P(^f|t) = 1, we get P(t) ~ 0.67
Therefore from (1),
     P(f|t) ~ 0.4
```

**Problem 3.** Work out the exercise described on slide 27 of lecture 5.

**Answer.** No difference

**Problem 4.** Work out the exercise described on slide 8 of lecture 6.

**Answer.** For $score = indegree + outdegree$: spamming is trivial. Simply include hundreds of links on the page. For $score = indegree$: spamming is still not too difficult. To increase the rank of a page, create a bunch of pages that link to it.

Note that one way to reduce the effect of spamming is to discount the weight of nepotistic links, i.e., links from within the same domain. Also note that Pagerank is much less susceptible to spamming than indegree.

**Problem 5.** Work out the exercise described on slide 14 of lecture 6.

**Answer.** Labelling the nodes 1, 2, 3 from left to right, we get the following transition probability matrix:

$$\begin{pmatrix} 0.0333 & 0.4833 & 0.4833 \\ 0.4833 & 0.0333 & 0.4833 \\ 0.0333 & 0.0333 & 0.9333 \end{pmatrix}$$

**Problem 6.** For the simple base set depicted in slide 26 of lecture 6, but with the $Alice \rightarrow MCI$ link removed, iterate the Hubs and Authorities algorithm 3 times.

**Answer.** For simplicity, the authority scores for the nodes on the left (which are 0 after the first iteration) and the hub scores for the nodes on the right (which are 0 after the first iteration) are not shown.