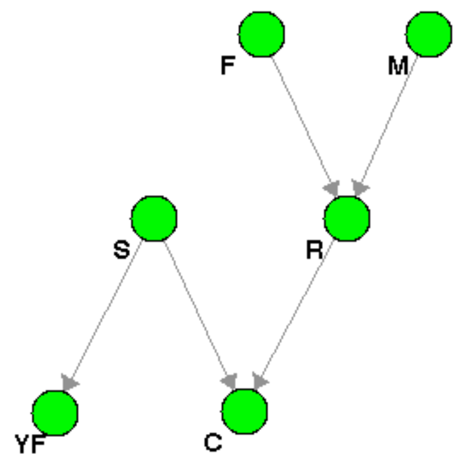
Problem 1:

The javaBayes network is given as:



1. The prior probability of cancer: P(C=1) = 0.53206 P(C=0) = 0.46794

Results of experiments:

```
probability (   "C" ) { //1 variable(s) and 2 values
 table
        0.5320600000000001 // p(true | evidence )
        0.46794;    // p(false | evidence );
}
```

2. The probability of smoking given cancer: P(S=1|C=1) = 0.40665, P(S=1|C=0) = 0.17874

Results of experiments:

```
 probability (   "S" ) { //1 variable(s) and 2 values
        table
                0.4066458670074804 // p(true | evidence )
                0.5933541329925196;        // p(false | evidence );
      }
 Posterior distribution:
 probability (   "smoking" ) { //1 variable(s) and 2 values
 table
        0.17874086421336066       // p(true | evidence )
        0.8212591357866393;        // p(false | evidence );
 }
```

3. The probability of smoking given cancer and radiation: P(S=1|C=1,R=1) = 0.3913 P(S=1|C=0,R=1) = 0.0968 P(S=1|C=1,R=0) = 0.5625 P(S=1|C=0,R=0) = 0.25

Results of experiments:

```
 probability (   "S" ) { //1 variable(s) and 2 values
        table
```

```
        0.39130434782608703     // p(true | evidence )
        0.6086956521739131;     // p(false | evidence );
    }
Posterior distribution:
probability (   "smoking" ) { //1 variable(s) and 2 values
table
    0.09677419354838711     // p(true | evidence )
    0.903225806451613;      // p(false | evidence );
}
Posterior distribution:
probability (   "smoking" ) { //1 variable(s) and 2 values
table
    0.25  // p(true | evidence )
    0.75;       // p(false | evidence );
}
Posterior distribution:
probability (   "smoking" ) { //1 variable(s) and 2 values
table
    0.5625000000000001 // p(true | evidence )
    0.4375;     // p(false | evidence );
}
```

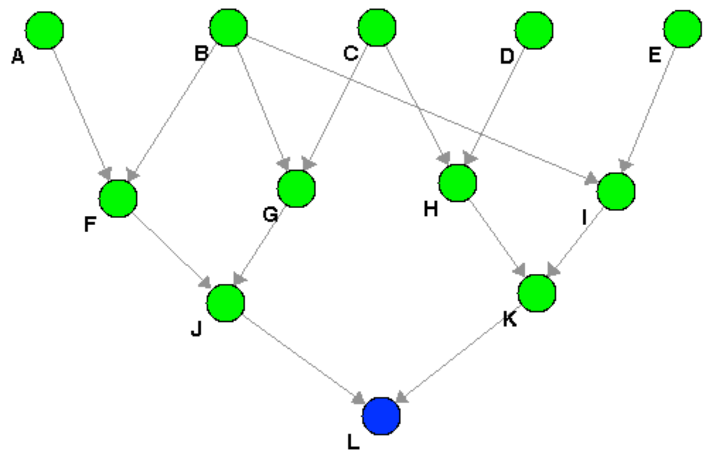
4. knowing cancer is present increases the probability of smoking. Knowing radiation reduces the probability of smoking, so the problem is saying cancer's contribution to the probability of smoking is decreased (from Piazza). Smoking and Radiation cause Cancer. Knowing Radiation is present decreases the probability of "Smoking caused cancer" .So that knowing radiation was present almost cancels the effect of knowing that cancer is present on the probability of smoking.
5. The markov blanket of yellow fingers: S(Smoking)
   the Markov blanket of node includes its parents, children and the other parents of all of its children
6. Solar flares and using the microwave are not independent given cancer, because the path between F and M are not blocked by C.
   Explanation: Node conditionally independent of its non-descendants given its parents.
   Node is conditionally independent of all other nodes in network given its parents, children, and children' s parents.
7. The probability of cancer if you never use a microwave: P(C=1$|\mathbf{M}=\mathbf{0}$) = 0.2554
   Results of experiments:

```
probability (   "C" ) { //1 variable(s) and 2 values
    table
        0.2554      // p(true | evidence )
        0.7446;     // p(false | evidence );
}
```

8. No, because not enough reasons causing cancer has been considered. The model is too naïve. The model could be improved by introduce more reasons, such as gender, age, race, environment, gene, and so on.

Problem 2:

The javaBayes network is given as:



P(A=pd|L=dd) = 0.02036

P(B=pd|L=dd) = 0.62413

P(C=pd|L=dd) = 0.38036

P(D=pd|L=dd) = 0.01880

P(E=pd|L=dd) = 0.01880

P(F=pd|L=dd) = 0.44898

P(G=pd|L=dd) = 0.80522

P(H=pd|L=dd) = 0.38575

P(I=pd|L=dd) = 0.62549

P(J=pd|L=dd) = 1.0

P(K=pd|L=dd) = 1.0

<u>Explanations:</u>

Since L is dd, J, K must both be disease carriers. Knowing J,K are pd, then the second generation has high probability of being pd. Since B is the parent of three of the second generation, it has high probabiltiy of being pd. And C is the parent of two, so it has second highest probability of being pd. While A, D and E are the parent of only one of the second generation, so they will have small probability of being pd comparing to B and C. Since G is the child of both B and C, so it would have the highest probability of being pd.

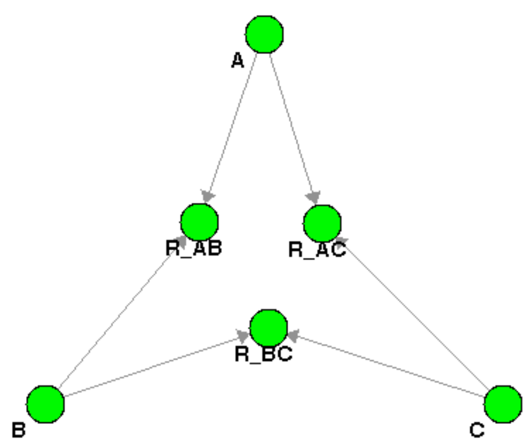
Results of experiments:

```
Posterior distribution:
probability (    "A" ) { //1 variable(s) and 2 values
       table
               0.0203626843919584 // p(pd | evidence )
               0.9796373156080416;       // p(pp | evidence );
}
Posterior distribution:
probability (    "B" ) { //1 variable(s) and 2 values
       table
               0.6241264848708957 // p(pd | evidence )
               0.3758735151291043;       // p(pp | evidence );
}
Posterior distribution:
probability (    "C" ) { //1 variable(s) and 2 values
       table
               0.38035713725489867       // p(pd | evidence )
               0.6196428627451014;       // p(pp | evidence );
}
Posterior distribution:
probability (    "D" ) { //1 variable(s) and 2 values
       table
               0.018797572033354353      // p(pd | evidence )
               0.9812024279666457;       // p(pp | evidence );
}
Posterior distribution:
probability (    "E" ) { //1 variable(s) and 2 values
       table
               0.01800913197300494       // p(pd | evidence )
               0.9819908680269951;       // p(pp | evidence );
}
Posterior distribution:
probability (    "F" ) { //1 variable(s) and 2 values
       table
               0.44898047861121004       // p(pd | evidence )
               0.55101952138879;   // p(pp | evidence );
}
Posterior distribution:
probability (    "G" ) { //1 variable(s) and 2 values
       table
               0.8052217172230584 // p(pd | evidence )
               0.19477828277694173;      // p(pp | evidence );
}
Posterior distribution:
probability (    "H" ) { //1 variable(s) and 2 values
       table
```

```
        0.38575306010385696     // p(pd | evidence )
        0.6142469398961431;     // p(pp | evidence );
}
Posterior distribution:
probability (   "I" ) { //1 variable(s) and 2 values
    table
        0.6254948280717493 // p(pd | evidence )
        0.3745051719282508;     // p(pp | evidence );
}
Posterior distribution:
probability (   "J" ) { //1 variable(s) and 2 values
    table
        1.0   // p(pd | evidence )
        0.0;  // p(pp | evidence );
}
Posterior distribution:
probability (   "K" ) { //1 variable(s) and 2 values
    table
        1.0   // p(pd | evidence )
        0.0;  // p(pp | evidence );
}
```

Problem 3:

The javaBayes network is given as:



1. In my Bayes net, A , B, C are three identical nodes, which stands for the quality of A, B and C, and has value 0, 1, 2, 3. R_AB, R_BC, R_AC are also three identical nodes, standing for the results between two teams. Because each team plays every other team once, the arcs I include is A->R_AB and B->R_AB, B->R_BC and C->R_BC, A->R_AC and C->R_AC. Arcs between A, B, C are not included, because the quality should be independent between teams.

2. For A, B, C, a reasonable prior probability table:

Q	P(Q)
0	0.25
1	0.25
2	0.25
3	0.25

For R_AB, R_AC, R_BC, a reasonable conditional probabiltiy table:

1,2 represents the two teams in one battle, can be A, B, C, R_12 has the probabilty of (1_wins, 2_wins, draw)

Q1 (1_wins, 2_wins, draw) Q2	0	1	2	3
0	(0.3,0.3,0,4)	(0.5,0.2,0.3)	(0.7,0.1,0.2)	(0.9,0,0.1)
1	(0.2,0.5,0.3)	(0.3,0.3,0,4)	(0.5,0.2,0.3)	(0.7,0.1,0.2)
2	(0.1,0.7,0.2)	(0.2,0.5,0.3)	(0.3,0.3,0,4)	(0.5,0.2,0.3)
3	(0,0.9,0.1)	(0.1,0.7,0.2)	(0.2,0.5,0.3)	(0.3,0.3,0,4)

3. The posterior distribution over the game between B and C, given no evidence: P(B_win_C) = 0.3625, P(C_win_B) = 0.3625, P(draw_BC) = 0.2750.

   The result makes sense because B, C are identical, so the probability for either one to win should be identical.

```
Posterior distribution:
probability (    "R_BC" ) { //1 variable(s) and 3 values
        table
                0.36249999999999993       // p(B_win | evidence )
                0.36250000000000004       // p(C_win | evidence )
                0.27499999999999997;      // p(draw | evidence );
}
```

4. The updated posterior distribution over the outcome of the game between B and C, given A beats B:

   P(B_win_C | A_win_B) = 0.2853, P(C_win_B | A_win_B) = 0.4405, P(draw_BC | A_win_B) = 0.2741.

   The results make sense because comparing to the results of the last question, the probability of B wining C decreases (since A beats B, B has high probability of having low quality, so the probability of B wins C should be decreased), and the probability of C wining B increases.

```
Posterior distribution:
probability (    "R_BC" ) { //1 variable(s) and 3 values
        table
                0.2853448275862069 // p(B_win | evidence )
                0.44051724137931036       // p(C_win | evidence )
                0.2741379310344828;       // p(draw | evidence );
}
```

5. If A beats B but draws against C, the updated posterior distribution over the outcome of the game between B and C:

   P(B_win_C | A_win_B, draw_AC) = 0.2616, P(C_win_B | A_win_B, draw_AC) = 0.4654, P(draw_BC | A_win_B, draw_AC) = 0.2730.

   The change makes sense. Since A wins B and A draws against C, C has higher probability of beating B. So the probability of B winning decreases, and C winning increases.

```
probability (    "R_BC" ) { //1 variable(s) and 3 values
        table
                0.2616352201257862 // p(B_win | evidence )
                0.46540880503144655       // p(C_win | evidence )
                0.27295597484276735;      // p(draw | evidence );
}
```

6. If A beats B but draws against C, the posterior distributions over the unobserved qualities of each team:

   P(A=0 | A_win_B, draw_AC) = 0.0943, P(A=1 | A_win_B, draw_AC) = 0.2075, P(A=2 | A_win_B, draw_AC) = 0.3208, P(A=3 | A_win_B, draw_AC) = 0.3774,

   P(B=0 | A_win_B, draw_AC) = 0.4151, P(B=1 | A_win_B, draw_AC) = 0.2925, P(B=2 | A_win_B, draw_AC) = 0.1887, P(B=3 | A_win_B, draw_AC) = 0.1038,

   P(C=0 | A_win_B, draw_AC) = 0.1808, P(C=1 | A_win_B, draw_AC) = 0.2531, P(C=2 | A_win_B, draw_AC) = 0.2909, P(C=3 | A_win_B, draw_AC) = 0.2752.

   The distributions make sense. Because A wins B, so A has high probability of having high quality, B has high probability of having low quality. And A draws against C, so that A and C has high probability of having the same quality, so that C also has higher probabiltity of having high quality than B.

```
Posterior distribution:
probability (    "A" ) { //1 variable(s) and 4 values
```

```
        table
                0.09433962264150943       // p(0 | evidence )
                0.2075471698113208 // p(1 | evidence )
                0.3207547169811321 // p(2 | evidence )
                0.37735849056603776;      // p(3 | evidence );
}
Posterior distribution:
probability (    "B" ) { //1 variable(s) and 4 values
        table
                0.4150943396226416 // p(0 | evidence )
                0.2924528301886793 // p(1 | evidence )
                0.18867924528301888       // p(2 | evidence )
                0.1037735849056604;       // p(3 | evidence );
}
Posterior distribution:
probability (    "C" ) { //1 variable(s) and 4 values
        table
                0.18081761006289312       // p(0 | evidence )
                0.2531446540880503 // p(1 | evidence )
                0.2908805031446541 // p(2 | evidence )
                0.27515723270440257;      // p(3 | evidence );
```

7. If A beats B (and the result of the game against C is unknown), the posterior distribution over C's quality:

P(C=0 | A_win_B) = 0.25, P(C=1 | A_win_B) = 0.25, P(C=2 | A_win_B) = 0.25, P(C=3 | A_win_B) = 0.25.

Because C’ s markov blanket does not include R_AB, so they are conditionally independent. Therefore the prior knowledge of R_AB does not influence the posterior distribution over C’ s quality.

8. If A is known to have quality 1, the posterior distribution over the outcome of the game between A and B:

   P(A_win_B | A = 0) = 0.275, P(B_win_A | A = 0) = 0.425, P(draw_AB | A = 0) = 0.3. This distribution makes sense. Comparing to question 3, the probability of A winning decreases, because it has a relatively low quality. And the probablity of B winning increases.

```
Posterior distribution:
probability (    "R_AB" ) { //1 variable(s) and 3 values
        table
                0.275// p(A_win | evidence )
                0.425// p(B_win | evidence )
                0.3;   // p(draw | evidence );
}
```