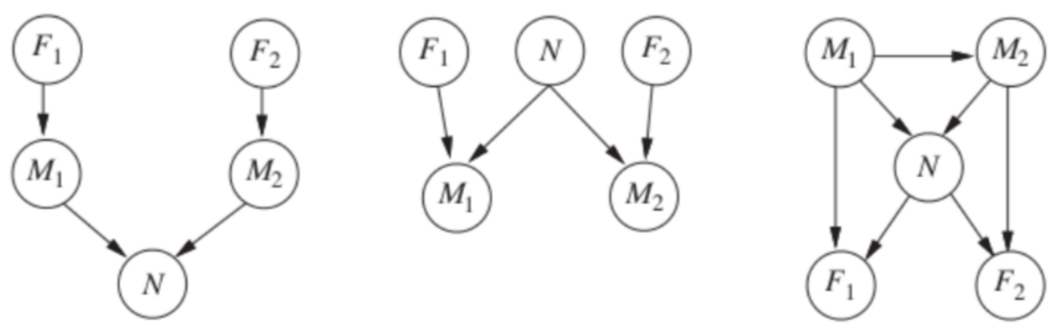
Name:	ID:

1. (14 points) Two astronomers are making measurements M_1 and M_2 of the number of stars N in a small region of sky using telescopes. The measurements are noisy, so there is a probability of e that the measurements are off by ±1. Besides this, the telescope might be out of focus with probability f (F_i = true means the i-th astronomers telescope is out of focus, i = 1,2). If it is out of focus, then the measurement will undercount N by at least 3 (or if N < 3, then M_i will just be 0). Consider the following.



(i) (ii) (iii)

a. (2 points) Which is the best network? Explain. (Hint: Accuracy and efficiency)

Best network: ii
Reason:
(i) suggests N (number of stars) is conditionally independent of Fi (telescope out of focus or not) given measurements Mi , which is incorrect from our description.
(iii) indicates some more information not described (the arrow from M1 to M2)

b. (9 points) Assuming using network (ii), and N ∈ {1,2,3} and M_1 ∈ {0,1,2,3,4}, write out the conditional distribution for $P(M_1|N)$ in terms of e and f . (Fill the blanks and equation)

We compute P(M1|N) by summing over (marginalizing) F1 (product rule and total probability).
P(M1|N) = P(M1|N, F1)P(F1|N) + P(M1|N, ¬F1)P(¬F1|N) = P(M1|N, F1)P(F1) + P(M1|N, ¬F1)P(¬F1)

$P(M_1 \| N)$	$N = 1$	$N = 2$	$N = 3$
$M_1 = 0$	f+e(1-f)	f	f
$M_1 = 1$	(1-2e)(1-f)	e(1-f)	0
$M_1 = 2$	e(1-f)	(1-2e)(1-f)	e(1-f)
$M_1 = 3$	0	e(1-f)	(1-2e)(1-f)
$M_1 = 4$	0	0	e(1-f)

P(M1|N) = P(M1, N)/P(N)
P(M1, N) = P(M1, N, F1) + P(M1, N, !F1)
= P(M1|N, F1)P(F1|N)P(N) + P(M1|N, !F1)P(!F1|N)P(N)
P(M1|N) = (P(M1|N, F1)P(F1|N)P(N) + P(M1|N, !F1)P(!F1|N)P(N))/P(N)
= P(M1|N, F1)P(F1|N) + P(M1|N, !F1)P(!F1|N)
= P(M1|N, F1)P(F1) + P(M1|N, !F1)P(!F1) F1 and N are independent

P(M1, N, F1) = P(M1|N, F1)P(N, F1) = P(M1|N, F1)P(F1|N)P(N)
= P(M1|N, F1)P(N|F1)P(F1)

M1=2, N=2
P(M1|N) = P(M1|N, F1)P(F1) + P(M1|N, !F1)P(!F1)
P(F1) = f
P(M1|N, F1) = 0
P(!F1) = 1-f
P(M1|N, !F1) = 1-2e
P(M1-N=1) = e P(M1-N=-1) = e P(M1=N) = 1-2e

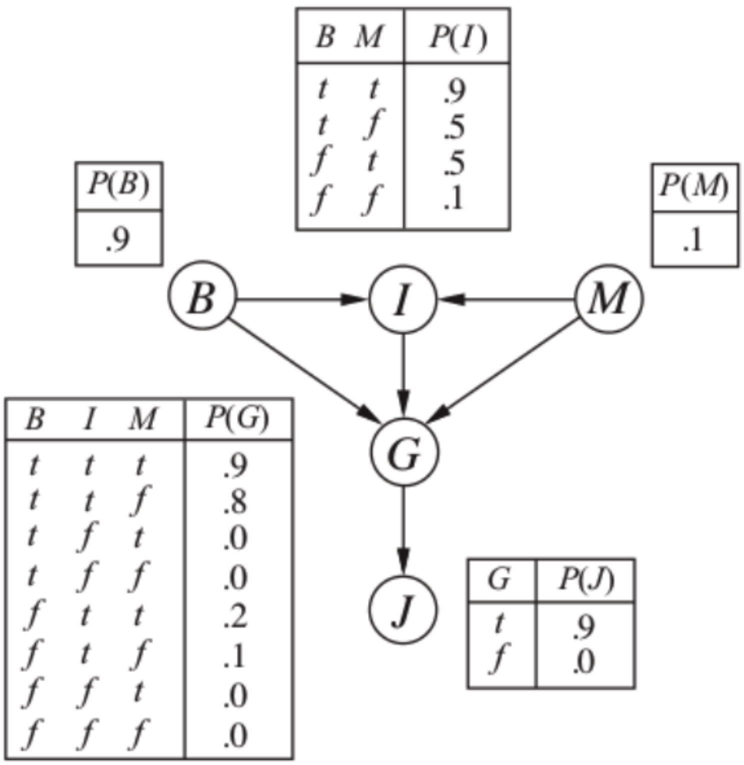
If F1=true, N=3, M1=0
P(M1|N=1, F1=true) = 1

c. (3 points) Suppose M_1 = 1 and M_2 = 3, and assume no prior constraint on N. What are the possible numbers that N can be? Why? (Hint: fix M1 = 1, M2 = 3, try each possibility of F1, F2 and get all possible values of N.)

start from N = 0, 1, . . . , and for each value, see whether M1 = 1 and M2 = 3 are possible with all possibilities of F1, F2

the solution should be N = 2, 4 or any number ≥ 6.

2. (6 points) Consider this figure for the next problem.



a. (3 points) Which of the following (if any) are asserted by the network structure (ignoring the conditional probability tables)? (Hint: Write "Yes" or "No" following each equation)

$P(B, I, M) = P(B)P(I)P(M)$; No

$P(J, G) = P(J|G, I)$; No

$P(M|G, B, I) = P(M|G, B, I, J)$; Yes
Markov blanket: a node is conditionally independent of all other nodes in the network, given its parents, children, and children's parents

b. (3 points) Calculate the value of $P(b, i, m, \neg g, j)$. (Fill the blanks and compute the result)

$$P(b, i, m, \neg g, j) \; = \; P(\qquad)P(m)P(i|b, m)P(\qquad\qquad)P(j|\neg g) \; =$$

P(b,i,m,¬g,j) = P(b)P(m)P(i|b,m)P(¬g|b,i,m)P(j|¬g) = 0.9×0.1×0.9×(1−0.9)×0.0
=0

Global semantics: P(X2, X1) joint probability
Marginalization: P(X1) = sum over X2, P(X1, X2)
Product: P(X2, X1) = P(X2|X1)P(X1) two events/semantics
P(X2|X1) = P(X2, X1)/P(X1)
P(x1, ..., xn) = P(xn |xn−1,...,x1)P(xn−1,...,x1)   n events/semantics

Goal: P(x1,...,xn) = P(xn |xn−1,...,x1)P(xn−1 |xn−2,...,x1) ⋯ P(x2 |x1)P(x1)

P(X2, X1) = P(X2|X1)P(X1) True
P(X3, X2, X1) = P(X3, X2 | X1)P(X1)
P(X3, X2, X1) = P(X3| X2, X1)P(X2, X1) True
P(X3, X2, X1) = P(X3| X2, X1)P(X2|X1)P(X1)
P(X4, X3, X2, X1) = P(X4| X3, X2, X1)P(X3| X2, X1)P(X2|X1)P(X1)
…
P(xn, …, x1) = P(xn | xn-1, …, x1)P(xn−1| xn-2 ..., x1)P(xn−2 | xn-3...,x1)P(xn-3,...,x1)
= P(xn |xn−1,...,x1)P(xn−1 |xn−2,...,x1) ⋯ P(x2 |x1)P(x1)

Independence:
If A, B are independent, P(A, B) = P(A)P(B); P(A|B) = P(A)
Generally, P(A, B) = (# of (A=True) & (B=True)) / ((# of (A=True) & (B=True)) + (# of (A=True) & (B=False)) + (# of (A=False) & (B=True)) + (# of (A=False) & (B=False)))
If independent: P(A, B) = (# of (A=True) / # all experiments (A=True or A=False)) * (# of (B=True) / # all experiments (B=True or B=False))
P(A| B) = (# of (A=True) & (B=True)) / (# (B=True))

1b
P(M1, N) = P(M1, N, F1) + P(M1, N, !F1)
#T times experiments
P(M1, N) = # (M1=m and N=n) / #T
= (# (M1=m and N=n and F1=true) + # (M1=m and N=n and F1=false)) / #T
= # (M1=m and N=n and F1=true) / #T + # (M1=m and N=n and F1=false) / #T
= P(M1, N, F1) + P(M1, N, !F1)
= P(M1|N, F1)P(F1|N)P(N) + P(M1|N, !F1)P(!F1|N)P(N)

$P(M1 \mid N) = P(M1, N) / P(N)$
F1 and N independent, $P(F1|N) = P(F1)$

#T
P(A) = #(A is true) / #T
P(M1-N==1) = e
P(M1-N==-1) = e

1c
$P(N|M1=1, M2=3)$
Valid N=n means $P(N=n|M1=1, M2=3) > 0$