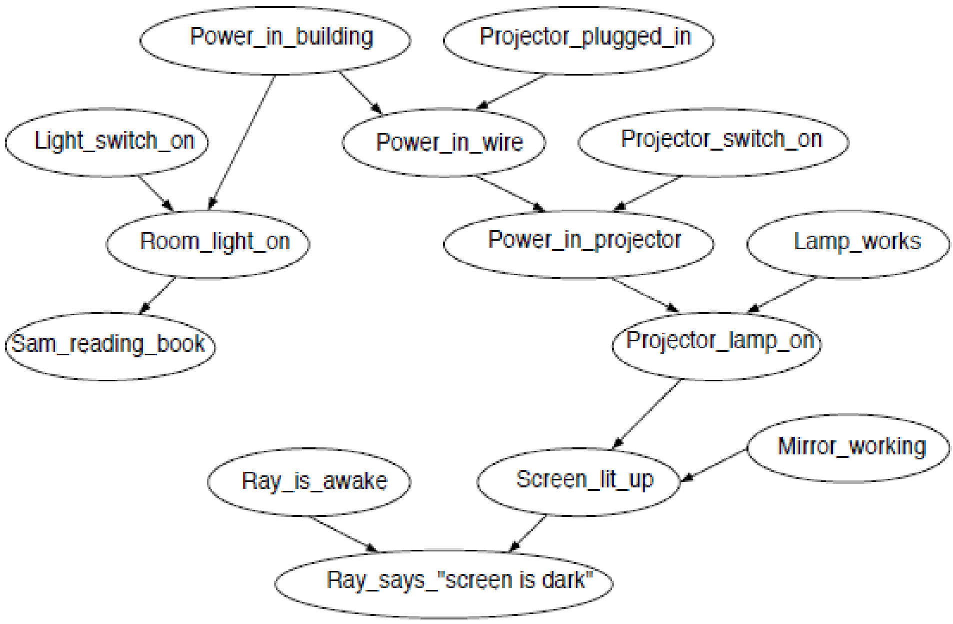
Bayesian Networks:

1. Consider the Bayesian network structure in the figure below. Using d-separation, answer the following questions. For the first three questions, either give an unblocked path between the two variables in the question or explain why all paths between the two variables in the question are blocked. (Hint: how many paths are there between to variables in this Bayesian network?) For the first three questions, also give an informal explanation of why d-separation gives the correct result, under the assumption that the names of the variables are meaningful, e.g., *Power_in _building* is a binary variable that represents whether there is power in the building where Sam may be reading, etc.

(a) Can knowledge of the value of *Projector_plugged_in* affect your belief in the value of Sam_reading_book?
(b) Can knowledge of *Screen_lit_up* affect your belief in *Sam_reading_book*?
(c) Can knowledge of *Projector_plugged_in* affect your belief in *Sam_reading_book* given that you have observed a value for *Screen_lit_up*?
(d) Which variables could have their probabilities changed if just *Lamp_works* were observed?
(e) Which variables could have their probabilities changed if just *Power_in_projector* were observed?



Answer

(a) No.
The (only) path between Projector_plugged_in and Sam_reading_book is blocked at the converging connection Power_in_wire, because there is no evidence on it or one of its descendents.
Whether the projector is plugged in is independent of whether Sam is reading.
(b) Yes.
The path between the two variables is not blocked, because it consists of serial connections and one diverging connection, and there is no evidence on any connections. It can affect our belief in whether there is power in the building.
(c) Yes.
The same path as in part (a) is now unblocked, because there is evidence on Screen_lit_up, which is a descendant of the converging connection Power_in_wire.
For example, knowing that the projector is not plugged in can explain why the screen is not lit up, which would then change our belief as to whether there was power in the building (which might have been a cause the screen was not lit up).

(d) Observing that Lamp_works can affect our belief in Projector_lamp on, Screen_lit_up and Ray_says_"screen is dark".
(e) Observing just Power in projector can affect our belief in all variables except Light_switch_on, Lamp_works, Mirror_working and Ray_is_awake.

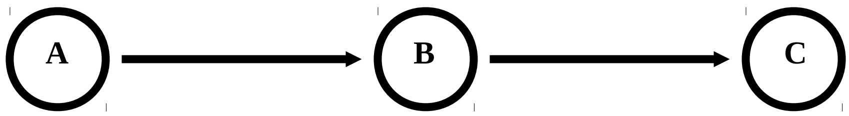
2. Two astronomers, in different parts of the world, make measurements M_1 and M_2 of the number N of stars in some small region of the sky, using their telescopes. Normally, there is a small probability of error of up to one star. Each telescope can also (with a slightly smaller probability) be badly out of focus (events F_1 and F_2), in which case the astronomer will underscore the count by three or more stars. Draw a Bayesian network structure that represents this story.

3.

P(A)
0.6

A	P(B|A)
T	0.8
F	0.3

B	P(C|B)
T	0.2
F	0.9



A. Given the above Bayesian network, compute P((B=true) AND (C=false)).
B Compute P(B)

3. You have a Bayesian network of N nodes. Each node corresponds to a Boolean random variable. Each node has a maximum of 3 parents. How many numbers would you need to specify at most, in order to fully specify the probability distribution modeled by this Bayesian network? In other words, what is the maximum number of values you need to store (for the entire network) in order to full specify the probability table for each node? Justify your answer.

4. After winning a race, an Olympic runner is tested for the presence of steroids. The test comes up positive, and the athlete is accused of doping. Suppose, it is known that 5% of all victorious Olympic runners do use performance enhancing drugs. For this particular test, the probability of a positive finding given that drugs are used is 95%. The probability of a false positive is 2%. What is the probability that the athlete did in fact use steroids, given the positive outcome of the test.

5. Draw the Bayesian networks that correspond to the following factorization of the joint probability distribution $P(A,B,C,D)$.

(i) $P(B|A,C)P(A)P(C|D)P(D)$
(ii) $P(A|B)P(C|B)P(B)P(D)$
(iii) $P(D|C)P(C|B)P(B|A)P(A)$
(iv) $P(B|A)P(A)P(C|D)P(D)$
(v) $P(A)P(B)P(C)P(D)$