Question 1.

Consider modelling a spam filter by means of a joint probability distribution $P(Y, X_1, \ldots, X_n)$ such that

$$Y = \begin{cases} 1 \text{ if the message is a spam,} \\ 0 \text{ otherwise.} \end{cases} \qquad X_i = \begin{cases} 1 \text{ if the message contains the } i\text{-th English word,} \\ 0 \text{ otherwise.} \end{cases}$$

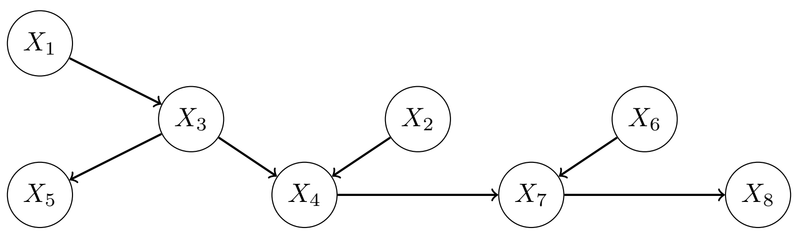
a) How many parameters do we need to store such distribution exactly? Do you have an estimate for n?

b) What problems may we encounter when trying to store the distribution exactly?

c) Do you have any ideas on how to improve the efficiency of the representation?

Question 2.

Consider the network (graph) below:



Decide the validity of the following statements:

a) $X_1 \perp\!\!\!\perp X_7 \mid X_3$

b) $X_1, X_5 \perp\!\!\!\perp X_6 \mid X_8$

c) $X_4 \perp\!\!\!\perp X_5 \mid X_1$

d) $X_1 \perp\!\!\!\perp X_2 \mid X_8$

e) $X_2 \perp\!\!\!\perp X_6$

f) $X_1 \perp\!\!\!\perp X_2, X_5$

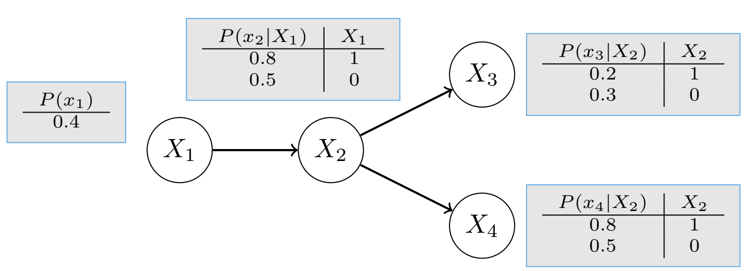
Question 3.

Consider the network below and compute

a) the marginal probability $P(X_3 = 0) = P(\neg x_3)$,

b) the pairwise marginal probability $P(X_2 = 1, X_3 = 0) = P(x_2, \neg x_3)$,

c) the conditional probability distribution $P(X_1 \mid X_2 = 1, X_3 = 0) = P(X_1 | x_2, \neg x_3)$.



Question 4.

Construct a Bayesian network (without CPTs) based on the following paragraph:

> When a family leaves their house, they often turn on the outdoor light. However, they also turn on the light when they are expecting a guest. The family has a dog, and they put it in the backyard when no one is home. They also put the dog there if it has bowel troubles. If the dog is in the backyard, it can probably be heard barking, although that could also be other dogs.

How many parameters do we need to represent such network?

How many parameters did we save compared to modelling the full joint distribution directly?