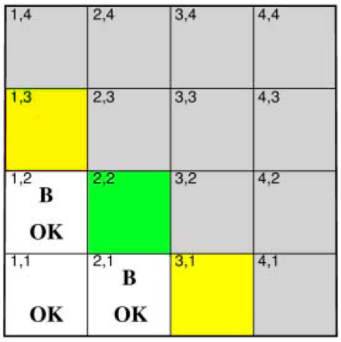
AI-III Exercises

1. **Wumpus World**

   As in lecture IV.2 consider the Wumpus world shown below.



   Squares directly adjacent to pits are breezy. We know there is nothing in $(1,1)$ and there are breezes in $(1,2)$ and $(2,1)$. Assume that a single field contains a pit with probability 0.2 and the pits are placed independently. What is the probability that there is a pit in $(2,2)$?

2. **Bayesian Networks**

   Consider the problem of detecting whether an e-mail contains a virus. We use four Boolean random variables to model this problem:

   V : indicates whether the e-mail contains a virus (true) or not (false);
   A : indicates whether the e-mail has an attachment (true) or not (false);
   L : indicates whether the e-mail is long (true) or not (false);
   K : indicates whether the sender is known to the receiver (true) or not (false).

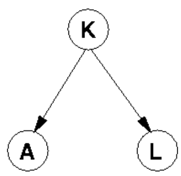
   (a) If we use a naive Bayesian approach, we make the assumption $\mathbf{P}(A, L, K|V) = \mathbf{P}(A|V)\,\mathbf{P}(L|V)\mathbf{P}(K|V)$. What does this assumption mean in terms of independence and/or conditional independence of any of the variables A, L, K,and V? Draw a Bayesian Network (without the conditional probability tables) that corresponds to this assumption. Justify its structure.

   (b) (i) Given an e-mail with A = true, L = false, K = false, compute the probability that it contains a virus using the probabilities listed in the table below under the assumption from Question 3(a).

   (ii) Can we fill in the probability table below with eight arbitrary real numbers? If so, say why, if not, list the conditions that the entries in the table must fulfil.

V	$P(A = \text{true}\|V)$	$P(L = \text{true}\|V)$	$P(K = \text{true}\|V)$	$P(V)$
false	0.2	0.8	0.8	0.8
true	0.9	0.2	0.1	0.2

   (c) Now assume someone constructed the following Bayesian Network for the variables A, L and K only:

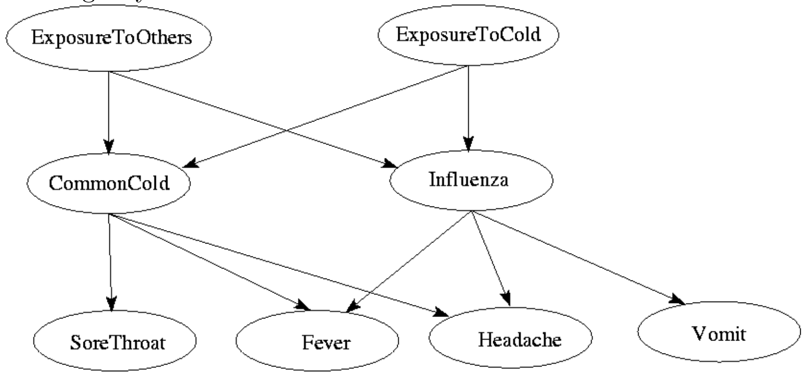


   Examining the lengths of 1000 e-mails whose senders were known to the receiver gave the following results. 200 of these e-mails were long. However, among the e-mails which had attachments, only 100 were long.

   Does this data support the assumptions made for the above Bayesian network? If not, suggest a suitable network structure. What implications, if any, has this result on the structure of the Bayesian network in Question 3(a)? Justify your answer.

3. **Diseases**

Consider the following Bayesian Belief Network:



(a) Which, if any, of the following are asserted by the structure of the above Bayesian Belief Network? Briefly justify your answer.

(1) $P(\texttt{CommonCold}, \texttt{ExposureToCold})$
$= P(\texttt{CommonCold})P(\texttt{ExposureToCold})$

(2) $P(\texttt{Vomit}) \neq P(\texttt{Vomit}|\texttt{ExposureToCold})$

(3) $P(\texttt{ExposureToCold}|\texttt{ExposureToOthers})$
$= P(\texttt{ExposureToCold})$

(4) $P(\texttt{SoreThroat}, \texttt{Vomit}|\texttt{CommonCold}, \texttt{Influenza})$
$= P(\texttt{SoreThroat}|\texttt{CommonCold})P(\texttt{Vomit}|\texttt{Influenza})$

(5) $P(\texttt{ExposureToOthers}|\texttt{ExposureToCold})$
$= P(\texttt{ExposureToOthers})$

(b) Based on the above Bayesian Belief Network derive an expression to compute the probability of the event `Influenza` given knowledge of whether the remaining events occurred.

(c) Based on the above Bayesian Belief Network, derive a formula to compute the probability of `Fever` $\wedge$ `Vomit` given knowledge of `CommonCold`, `Influenza` and `ExposureToCold`. What is the relation between the events `Fever` and `Vomit`?

(d) Derive Bayes' rule from the definition of conditional probability.

4. **Politics**

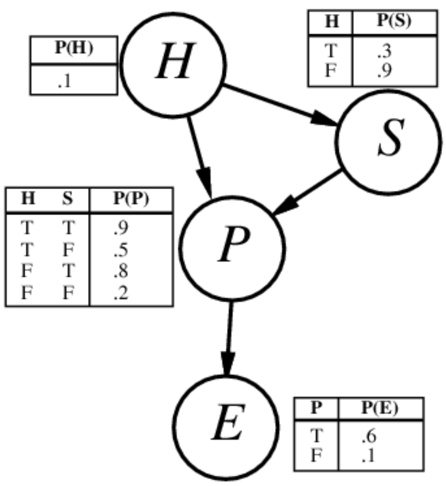
Consider the following simple belief network with boolean variables **H** = Honest, **S** = Slick, **P** = Popular, **E** = Elected.

(a) Which, if any, of the following are asserted by the network structure (ignoring the CPTs for now)? Note that any subset of these may be correct.

(i) $\mathbf{P}(\mathbf{H}, \mathbf{S}) = \mathbf{P}(\mathbf{H})\mathbf{P}(\mathbf{S})$

(ii) $\mathbf{P}(\mathbf{E}|\mathbf{P}, \mathbf{H}) = \mathbf{P}(\mathbf{E}|\mathbf{P})$

(iii) $\mathbf{P}(\mathbf{E}) \neq \mathbf{P}(\mathbf{E}|\mathbf{H})$



(b) Calculate the value of $P(h, s, \neg p, \neg e)$

(c) Calculate the probability that someone is elected given that they are honest.

(d) Suppose we want to add the variable **L** = LotsOfCampaignFunds to the network; describe, with justifications, all the changes you would make to the network.