Exercise: Removing leaves in BN20 networks

1. Consider the QMR network, where only some of the symptoms are observed. For example, in Figure 1(a), X_3 and X_5 are hidden. Show that we can safely remove all the hidden leaf nodes without affecting the posterior over the disease nodes, i.e., prove that we can compute $p(\mathbf{z}_{1:3}|x_1, x_2, x_4)$ using the network in Figure 1(b). This is called **barren node removal**, and can be applied to any DGM. (A barren node is a hidden leaf node.)

2. Now suppose we partition the leaves into three groups: on, off and unknown. Clearly we can remove the unknown leaves, since they are hidden and do not affect their parents. Show that we can analytically remove the leaves that are in the "off state", by absorbing their effect into the prior of the parents. (This trick only works for noisy-OR CPDs.)

Figure 1: (a) A QMR-style network with some hidden leaves. (b) Removing the barren nodes.