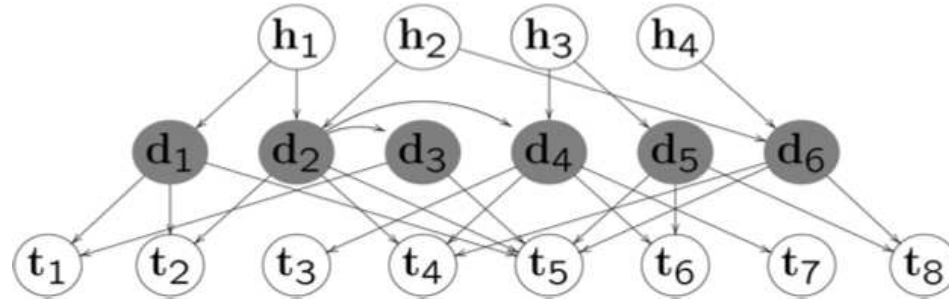


Simple Promedas implementation



$h = 0, 1$ denote conditions; $d = 0, 1$ denote diagnoses; t denote tests.

Ignore conditions and assume tests binary ($t = 0, 1$).

$$p(t_1, \dots, t_m, d_1, \dots, d_n) = p(d_1) \dots p(d_n) p(t_1 | d_{pa(1)}) \dots p(t_m | d_{pa(m)})$$

$p(d_j = 1)$ is the prevalence of the disease. $p(t_i | d_{pa(i)})$ is the probability of test outcome $t_i = 0, 1$ for a given configuration of the diagnoses. $d_{pa(i)}$ is the subset of diagnoses that affect test i . (For instance, $pa(1) = d_1, d_3$ in the figure.)

A simple choice of the function $p(t_i|d_{\text{pa}(i)})$ is:

$$p(t_i = 1|d_{\text{pa}(i)}) = \sigma\left(\sum_{j \in \text{pa}(i)} w_{ij}d_j + w_{i0}\right)$$

$$\sigma(x) = \frac{1}{2}(1 + \tanh(x))$$

$$p(t_i = 0|d_{\text{pa}(i)}) = 1 - p(t_i = 1|d_{\text{pa}(i)}) = \sigma\left(-\sum_{j \in \text{pa}(i)} w_{ij}d_j - w_{i0}\right)$$

Given some patient findings $t_{1:k} = t_1, \dots, t_k$, the probability of diagnose j is given by Bayes formula:

$$p(d_j | t_{1:k}) = \frac{p(d_j, t_{1:k})}{p(t_{1:k})}$$

$$\begin{aligned} p(d_j, t_{1:k}) &= \sum_{d_{1:n \setminus j}} \sum_{t_{k+1:m}} p(d_1) \dots p(d_n) p(t_1 | d_{\text{pa}(1)}) \dots p(t_m | d_{\text{pa}(m)}) \\ &= \sum_{d_{1:n \setminus j}} p(d_1) \dots p(d_n) p(t_1 | d_{\text{pa}(1)}) \dots p(t_k | d_{\text{pa}(k)}) \end{aligned}$$

The marginal computation uses a graph defined on the diagnoses that are parents of some test.

Example: patient case with one positive test $t_1 = 1$ and one negative test $t_2 = 0$
 t_1 has parents d_1, d_2 , t_2 has parents d_2, d_3 :

$$\begin{aligned}
 & p(d_1, d_2, d_3) \\
 = & \frac{p(d_1)p(d_2)p(d_3)\sigma(w_{11}d_1 + w_{12}d_2 + w_{10})\sigma(-w_{22}d_2 - w_{23}d_3 - w_{30})}{\sum_{d_1, d_2, d_3} p(d_1)p(d_2)p(d_3)\sigma(w_{11}d_1 + w_{12}d_2 + w_{10})\sigma(-w_{22}d_2 - w_{23}d_3 - w_{30})} \\
 = & \frac{1}{Z} \exp \left(\sum_j \log p(d_j) + \log \sigma(w_{11}d_1 + w_{12}d_2 + w_{10}) + \log \sigma(-w_{22}d_2 - w_{23}d_3 - w_{30}) \right) \\
 = & \frac{1}{Z} \exp \left(\sum_j \alpha_j d_j + \sum_{jj'} \alpha_{jj'} d_j d_{j'} \right)
 \end{aligned}$$

$$\alpha_j = \log \frac{p(d_j = 1)}{p(d_j = 0)} + \dots$$

act as external fields (magnetizations).

$\alpha_{jj'}$ depends on all w_{ij} . Can be of either sign.

The probability distribution is like an Ising model, but in general with interaction terms involving more than just two diagnoses.

Marginalization

For instance, the probability on d_1 given the patient findings $t_1 = 1$ and $t_2 = 0$

$$p(d_1) = \frac{1}{Z} \sum_{d_2, d_3, \dots} \exp \left(\sum_j \alpha_j d_j + \sum_{i=1}^k \log \sigma \left(\sum_{j \in pa(i)} w_{ij} d_j + w_{i0} \right) \right)$$

Different methods for marginalization:

- exact summation when the number of variables is relatively small
- Monte Carlo sampling (Metropolis Hasting or Gibbs sampling...)
- approximate analytical methods

We learned from the Ising model that Monte Carlo sampling is effective when

- couplings α (ie. w) are small, or
- couplings are all positive (no frustration)