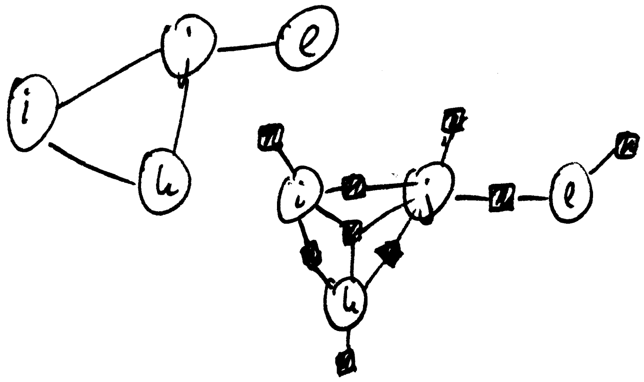
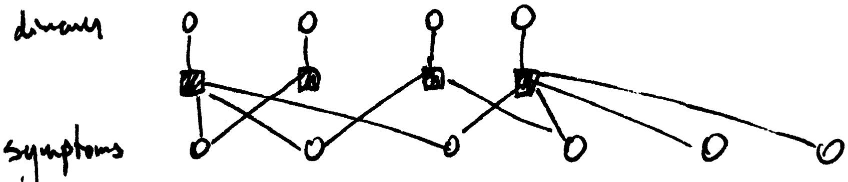
* Bayes networks
* Markov Random Fields — Global Markov Property
* Gibbs Random Field — dist factorises over cliques in the graph



* Factor Graphs V, F bipartite graphs

<u>Example</u> Disease-symptom network



Symptoms and disease variables are binary.

$$\psi_a(x_{\partial a}) = \begin{cases} 0 & \text{if } x_d = 1 \;\; x_{\partial a} \neq 1 \\ 1 & \text{otherwise} \end{cases}$$

∂a is the disease node associated to a.

$$f(x_V) = \frac{1}{Z_\psi} \prod_{a \in F} \psi_a(x_{\partial a})$$

Conditioning in a factor graph on a value x_O^* for the observable nodes $O \subseteq V$ is simple; we define the potentials

$$\phi_a(x_{\partial a}) = \psi_a(x_{\partial a}) \, \mathbb{1}(x_{\partial a \cap O} = x^*_{\partial a \cap O})$$

<u>Theorem</u> Hammersley-Clifford

Suppose $f(x_V) > 0$ for all $x_V \in \mathbb{X}_V$, satisfies the G.M.P. with respect to G. then X_V is also a G.R.F. w.r.t. G. Conversely, any G.R.F satisfies the G.M.P.

<u>Pf:</u> $$f(x_V) = \frac{1}{Z_\psi} \prod_{c \text{ a clique}} \psi_c(x_c)$$

Take a partition $V = V_1 \uplus V_2 \uplus A$ s.t. A separates V_1 & V_2, want $X_{V_1} \perp X_{V_2} \mid X_A$.
Since no clique contains both $v_1 \in V_1$ & $v_2 \in V_2$ then product splits into 2 factors one depending on x_{V_1}, the other on x_{V_2}.

For any $S \subseteq V$ define

$$\tilde{\psi}_S(x_S) = \prod_{U \subseteq S} f(x_U, 0_{V \setminus U})^{(-1)^{|S|-|U|}} > 0$$ where 0_V is an arbitrary configuration.

Claim

1) $f(x_V) = \prod_{S \subseteq V} \tilde{\psi}_S(x_S)$

2) If S is not a clique
$\tilde{\psi}_S(x_S) = 1$

1) Consider any $U \subset S$, how many times we have a factor $f(x_U, 0_{V\setminus U})$ with power 1 & -1 in the product $\prod_{S \subseteq V} \tilde{\psi}_S(x_S)$. There are $\binom{|V|-|U|}{k}$ sets S s.t. $U \subseteq S \subseteq V$ with $|S| - |U| = k$. Therefore, the total power of $f(x_U, 0_{V\setminus U})$ is

$$1 - \binom{|V|-|U|}{1} + \binom{|V|-|U|}{2} - \dots + (-1)^{|V|-|U|}\binom{|V|-|U|}{|V|-|U|} = (1-1)^{|V|-|U|} = 0$$

2) If S not a clique, it contains vertices a, b not connected in G.

$$\tilde{\psi}_S(x_S) = \prod_{U \subseteq S \setminus \{a,b\}} \Bigg[\underbrace{\frac{f(x_U, 0_{V\setminus U})}{f(x_{U\cup\{a\}}, 0_{V\setminus U \setminus \{a\}})}}_{(*)} \underbrace{\frac{f(x_{U\cup\{a,b\}}, 0_{V\setminus U\setminus\{a,b\}})}{f(x_{U\cup\{b\}}, 0_{V\setminus U\setminus\{b\}})}}_{(**)} \Bigg]^{\pm 1}$$

$$(*) = \frac{f(0_a \mid x_U, 0_{V\setminus U\setminus\{a\}})}{f(x_a \mid x_U, 0_{V\setminus U\setminus\{a\}})} \underset{\text{G.M.P.}}{=} \frac{f(0_a \mid x_{U\cup\{b\}}, 0_{V\setminus U\setminus\{a,b\}})}{f(x_a \mid x_{U\cup\{b\}}, 0_{V\setminus U\setminus\{a,b\}})} = \frac{1}{(**)}$$

X_a $\quad$ X_b $\quad$ ∎

Inference problems.

1) compute partition function Z_ψ
2) Sampling X_V
3) Compute conditional dist $f(x_A \mid x_B)$ for $A, B \subseteq V$
4) ~~Marginalization~~ Computing marginals $f(x_A)$ for $A \subseteq V$

Computational reductions

③ $\longrightarrow$ ④ : Special case, set $B = \emptyset$

④ $\longrightarrow$ ③ : Apply Bayes theorem $f(x_A \mid x_B) = \dfrac{f(x_{A\cup B})}{f(x_B)}$

④ $\longrightarrow$ ② : Order the variables in an arbitrary way $v_1, v_2, \dots, v_m$, sample

$$X_{v_1} \sim f_{X_{v_1}}$$

$$X_{v_2} \sim f_{X_{v_2} \mid X_{v_1}}$$

$$\vdots$$

$$X_{v_m} \sim f_{X_{v_m} \mid X_{v_1} \dots X_{v_{m-1}}}$$

② $\longrightarrow$ ④ : Sample realizations of $x_V^{(1)}, x_V^{(2)}, \dots, x_V^{(N)}$, estimate $f(x_A) \approx \frac{1}{N}\sum_{i=1}^{N} \mathbb{1}(x_A^{(i)} = x_A)$

① $\longrightarrow$ ④ : Let ϕ be the potentials for the reduced model conditioned on $X_A = x_A^*$

$$f(x_A^*) = \sum_{\substack{x_V \\ x_A = x_A^*}} \frac{1}{Z_\psi} \prod_{a\in F} \psi_a(x_{\partial a}) = \frac{1}{Z_\psi} \sum_{x_V} \prod_{a \in F} \phi_a(x_{\partial a}) = \frac{Z_\phi}{Z_\psi}.$$

④ $\longrightarrow$ ① : ...