Undirected Graphical Models
Inference in graphical models

A5: marked version due now

A6: due now

CP: out tonight (pick your model at 5pm)
Auditors too

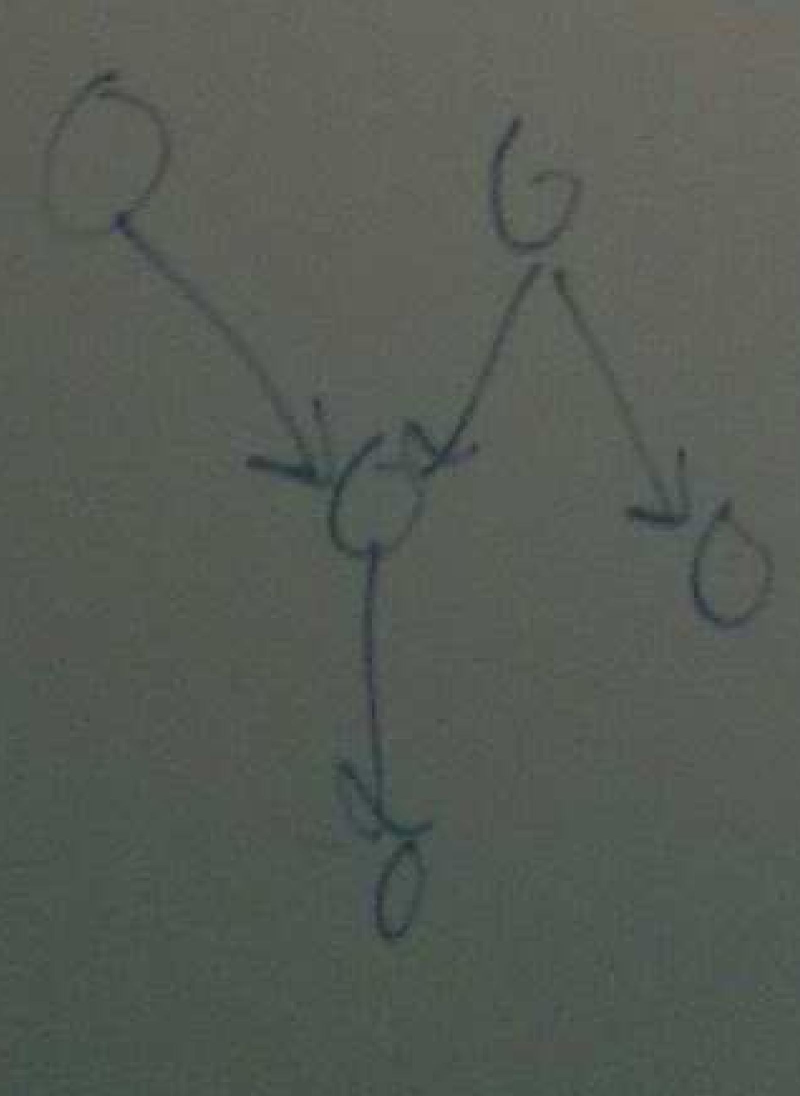
Midterm: Monday

A7: out Monday (N.B. → only top 6 assignments count)

Last time: DAG models

$$p(x) = \prod_{j=1}^{d} p(x_j \mid x_{\pi(j)})$$

w/ notation $x_j = (x_{:})_j$



Undirected Graphical Models

"Markov random fields"
"Markov networks"

Divide $\{x_1, x_2, \ldots, x_d\}$ into (possibly overlapping) subsets 'c'.

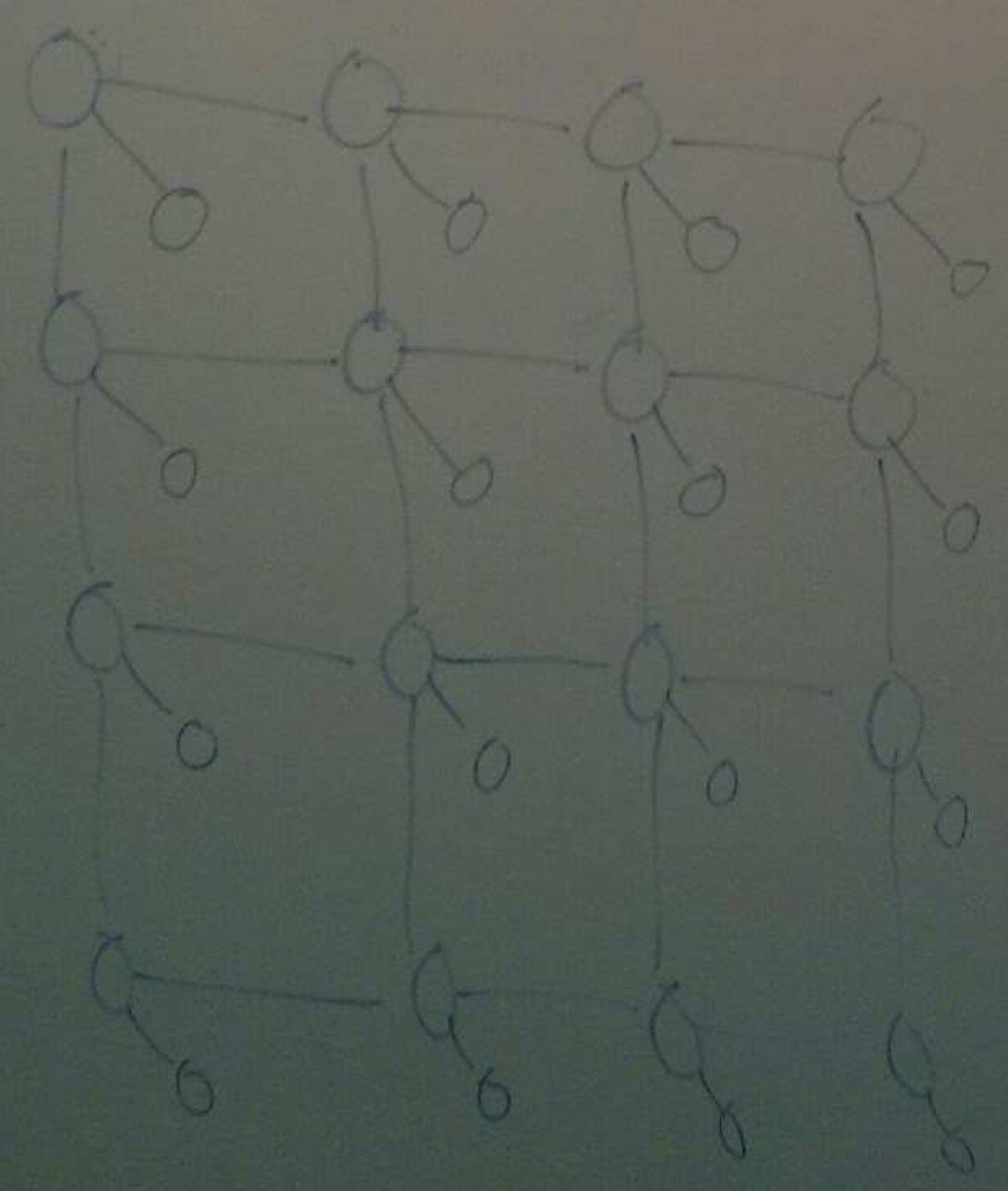
$$p(x) = \frac{1}{Z} \prod_{c \in C} \phi_c(x_c)$$

Graph:

V: variables x_j

E: edge $j-k$ if x_j and x_k appear in some subset 'c'.

Example: image segmentation



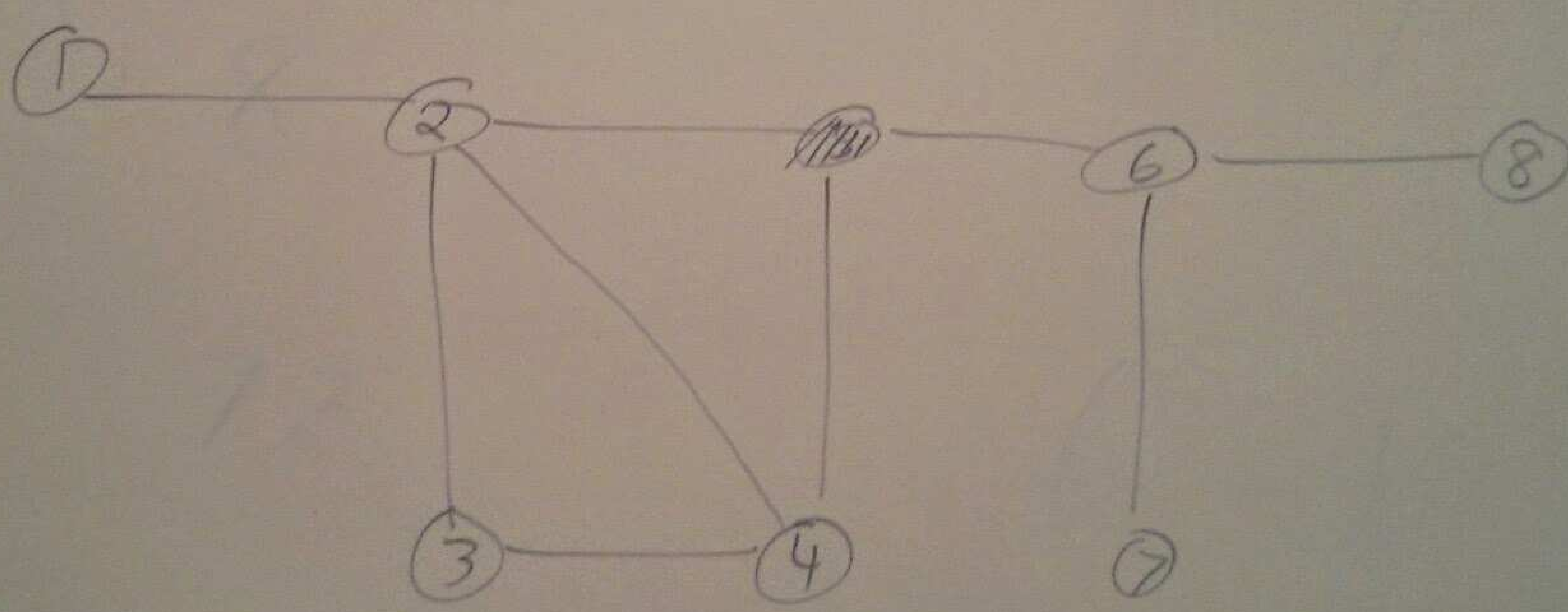
"Local Markov property":

$$x_j \perp x_{(1\,j)\setminus \underline{nei(j)}} \mid x_{nei(j)}$$

- determine if independence follows from using graph separation:

"  "

R^+



$$p(x) = \psi_1(x_1)\,\psi_{12}(x_{12})\,\psi_{23}(x_{23})\,\psi_{34}(x_{34})\,\psi_{25}\,\psi_{45}\,\psi_{56}\,\psi_{67}\,\psi_{68}\,\psi_{234}(x_{234})$$

$x_1 \not\perp x_2 \mid x_5$

$x_1 \perp x_6 \mid x_5$

Pairwise UGMs

$$p(x) = \frac{\prod_{j=1}^{d}\left[\psi_j(x_j)\prod_{(j,k)\in E}\psi_{jk}(x_j,x_k)\right]}{Z}$$

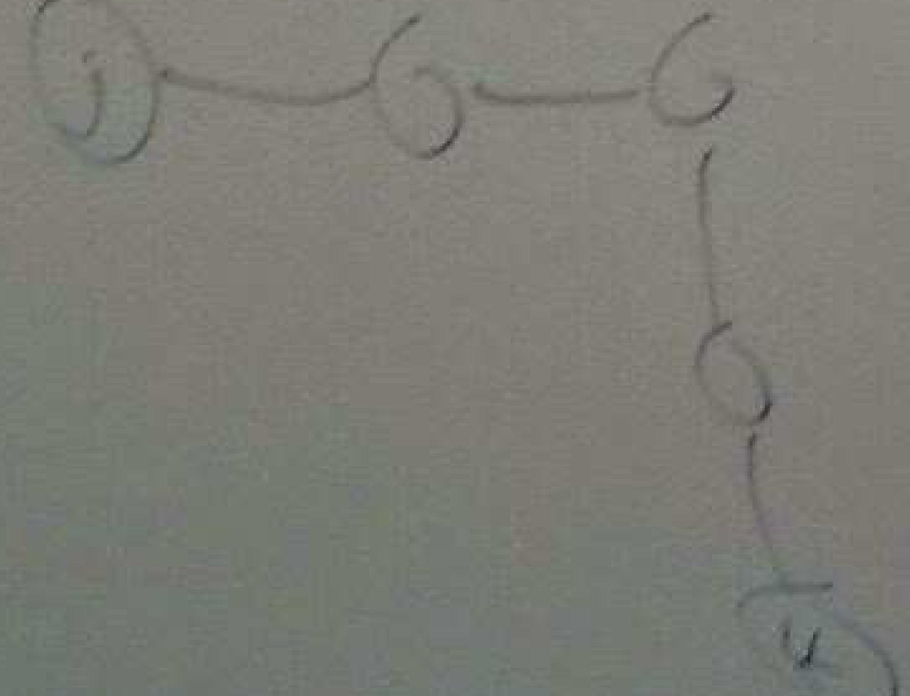
Gaussian Graphical Models

$x \sim N(\mu,\Sigma)$

$\Sigma_{jk}=0$

$\Downarrow$

x_j is not reachable from x_k



$(\Sigma^{-1})_{jk}=0$

$\Downarrow$

edge $x_j - x_k$ is missing

Ising models

$x_j \in \{-1,1\}$

$\psi_j(x_j) = \exp(-w_j x_j)$

$\psi_{jk}(x_j,x_k) = \exp(w_{jk} x_j x_k)$

$$p(x) = \frac{\exp\left(-\sum_j\left[w_j x_j - \sum_{(j,k)\in E} w_{jk} x_j x_k\right]\right)}{Z}$$

If $x_j \in \{1,2,\ldots,S\}$,

"Potts" model

Special case of "log linear models"

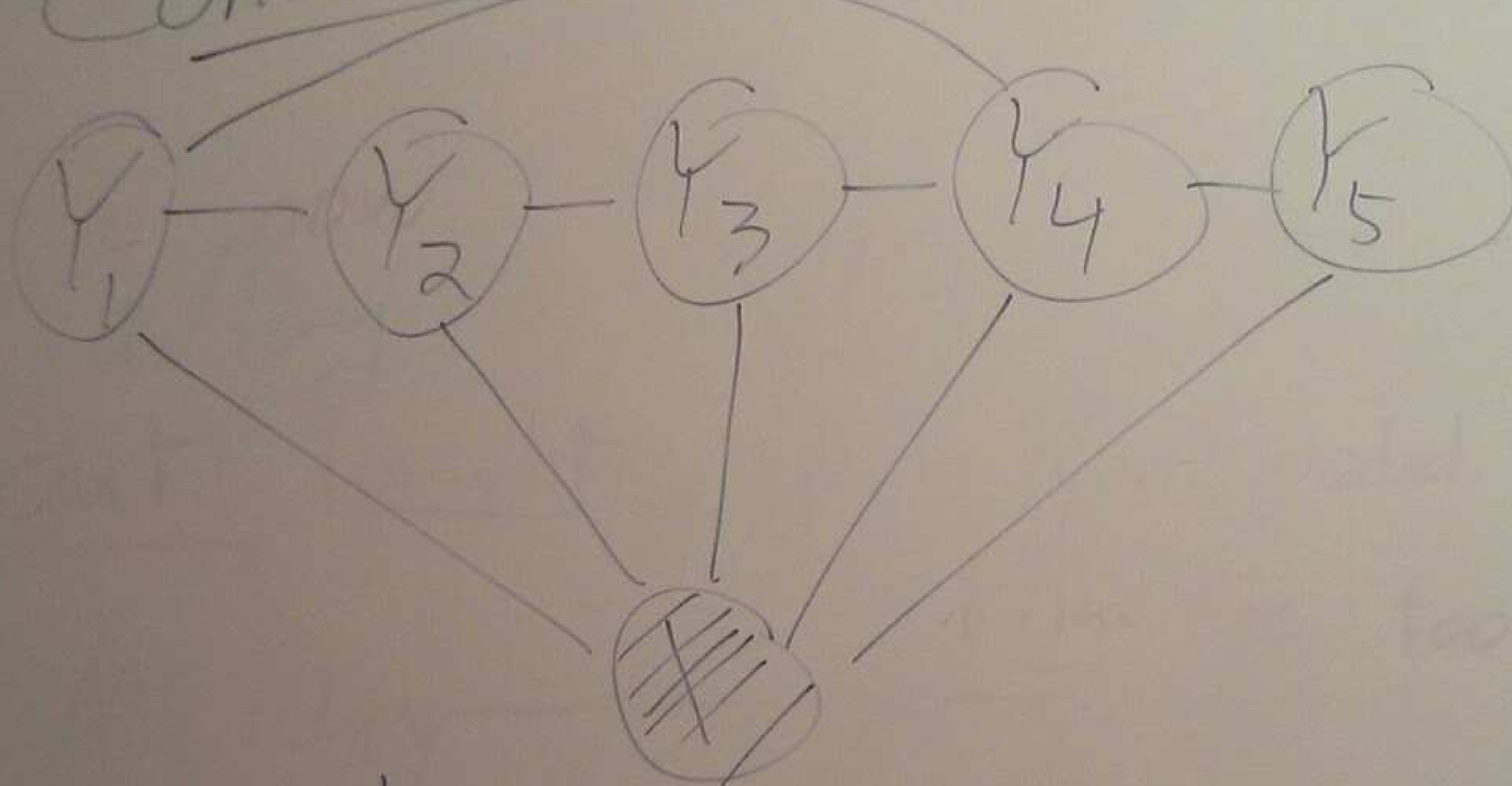
Ass

ndirected Graphical Models
nference in graphical models

n: $\{X, y\}$

t: $\hat{X} \to \hat{y}$

Conditional Random Field

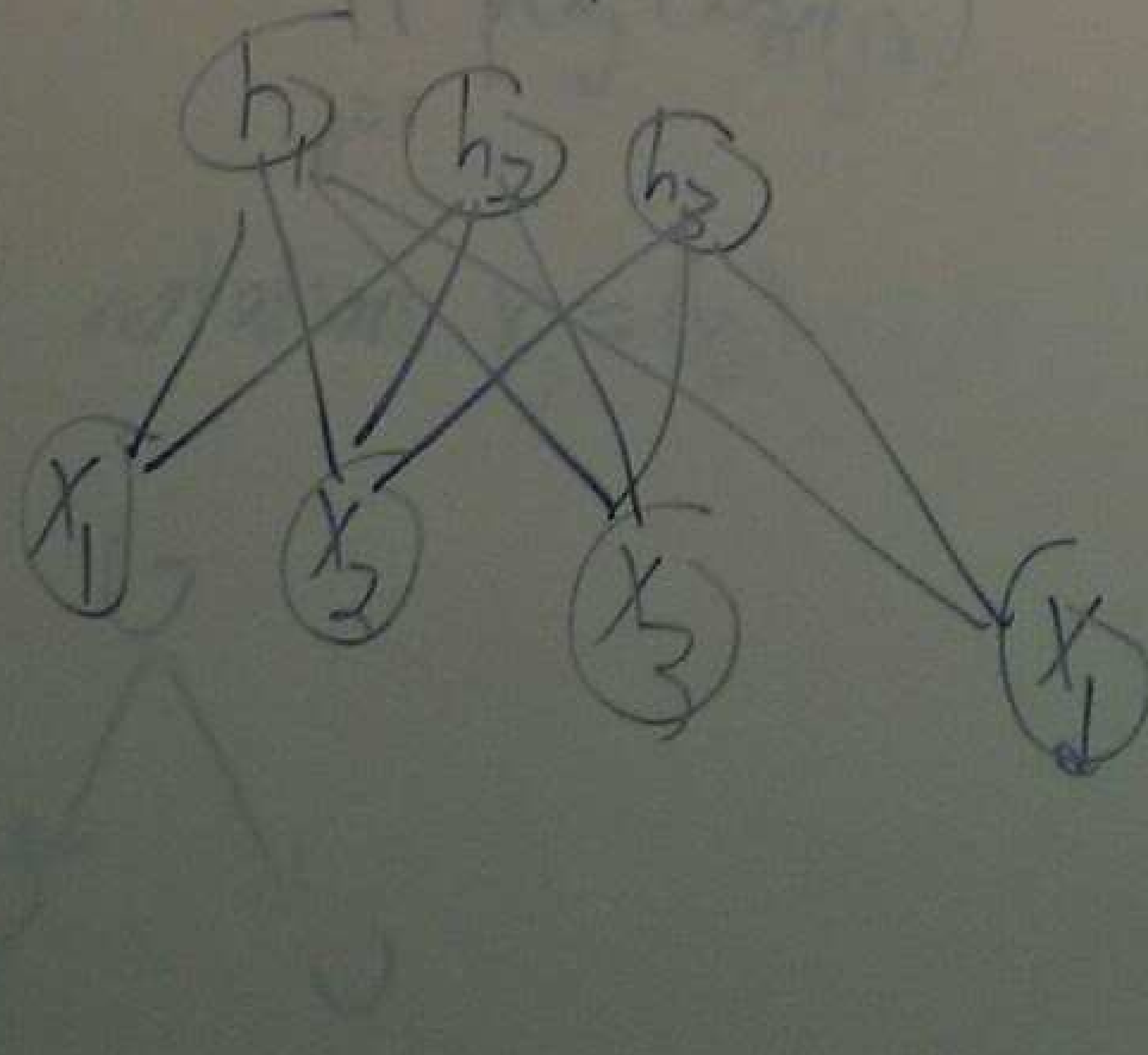


Logistic is special case.

Boltzmann machines:

Ising model w/ hidden variables.

"Restricted" BM:



Computation
(discrete)

compute:

$p(x) \rightarrow$ easy if you know Z

$p(x) \rightarrow$ #P-hard

$p(x_j) \rightarrow$ #P-hard

$p(x_j \mid x_k) \rightarrow$ #P-hard

$\max_x p(x) \longrightarrow$ NP-hard

We can solve these efficiently if 'G' is nice.

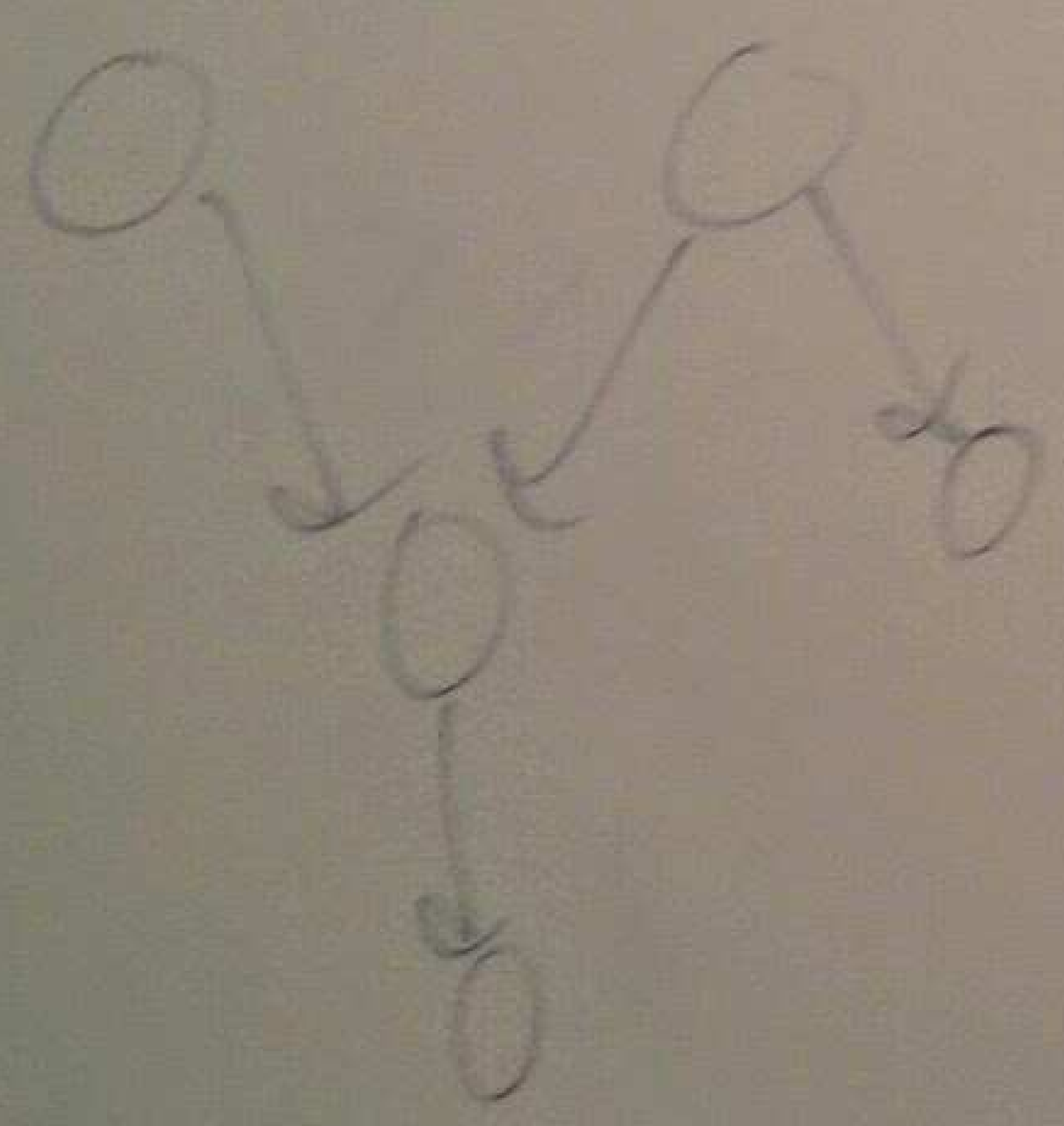
DAGs vs. UGMs

Easy ops DAGs: $p(x)$

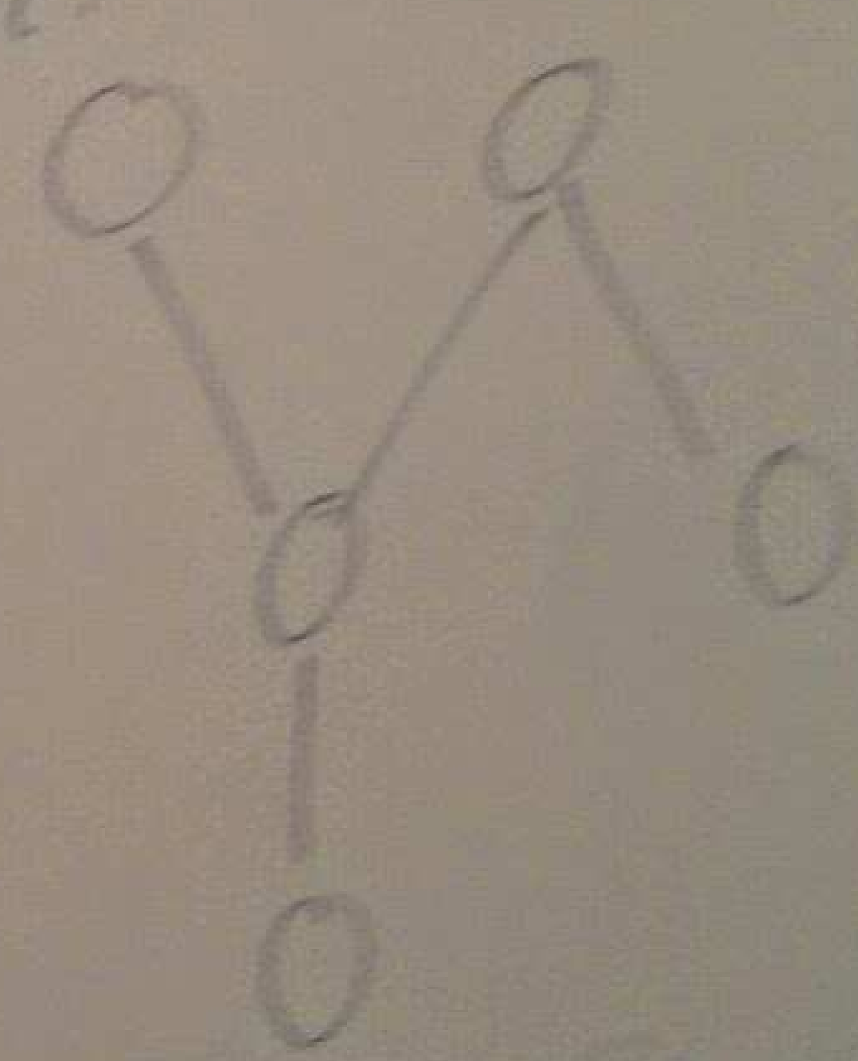
$p(x_j \mid x_{\pi_j})$

parameter estimation

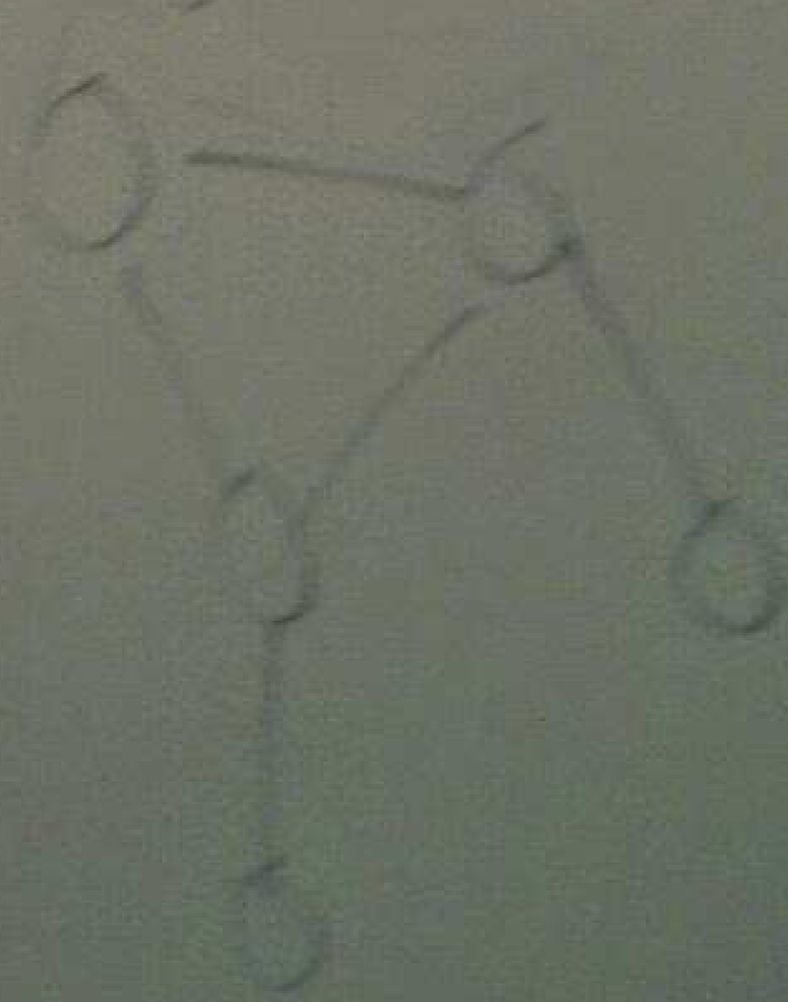
Everything else: as hard as UGMs
- convert to UGM and solve.



remove directions



"moralization"
"marry unwed parents"



Viterbi decoding

Consider chain-structured UGM $x_1 - x_2 - x_3 - \cdots - x_d$

$$p(x) = \frac{\psi_1(x_1)\prod_{j=1}^{d-1}\psi_j(x_j, x_{j+1})}{Z}$$

"Decoding": $\max_x p(x)$

"most likely sequence"

If $x_j \in \{1, 2, \ldots, s\}$, there s^d configurations.

But, can be solved in $O(ds^2)$ by "dynamic programming"

DP: strategy for solving discrete optimization, where solution is defined recursively

rents'

Ass

directed Graphical Models

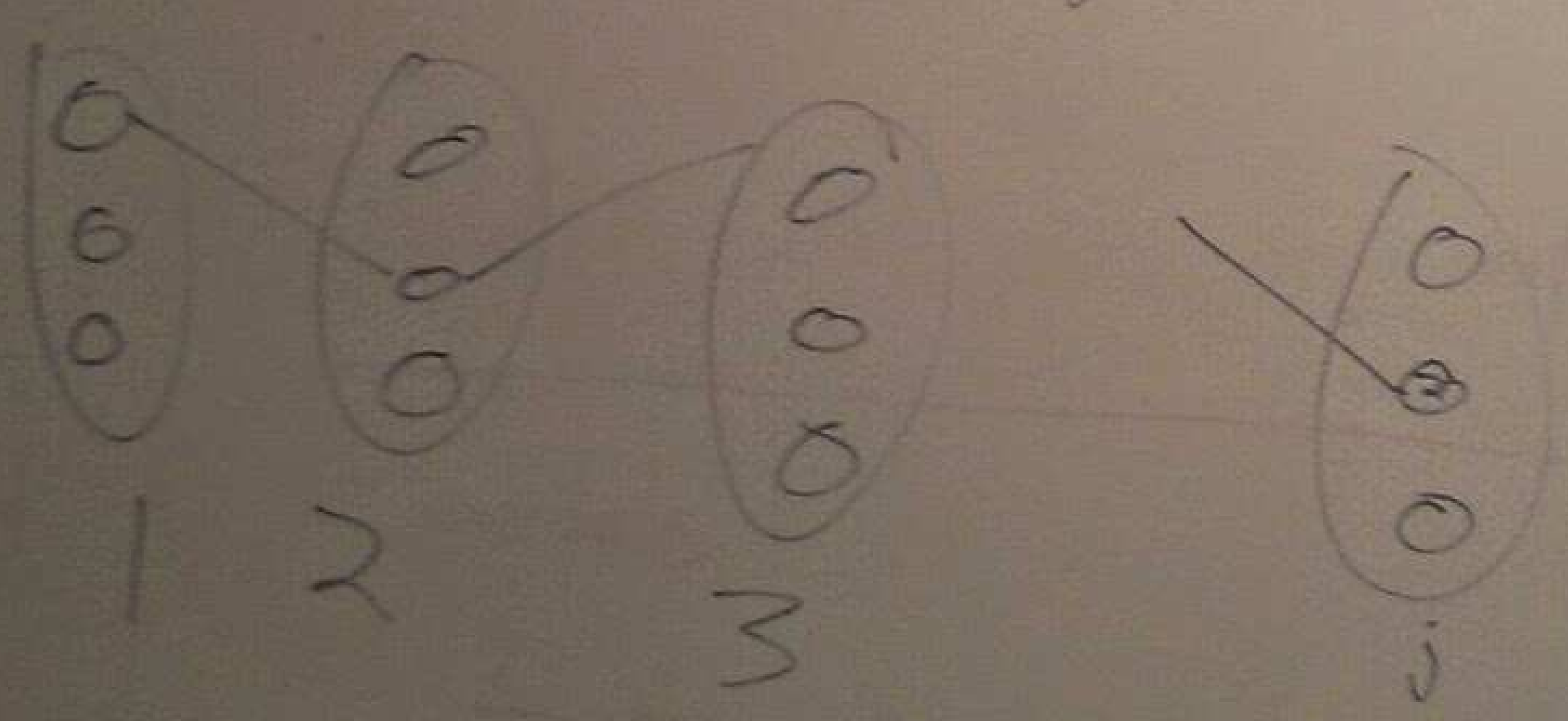
ference in graphical models

Let V_{jk} be the potential of

the most probable sequence

based $\psi_1(x_1), \psi_1(x_1, x_2), \psi_2(x_2, x_3), \ldots, \psi_{j-1}(x_{j-1}, x_j)$,

that ends with $x_j = k$.



Optimal solution: $\max_k V_{dk}$

$$V_{1k} = \psi_1(x_1 = k)$$

$$V_{jk} = \max_{k'} \{ \psi_{j-1}(x_{j-1} = k', x_j = k) V_{(j-1)k'} \}$$

$$V = s\left\{\begin{array}{l}\psi_1(x_1=1) \xrightarrow{\max} V_{11}\\ \psi_2(x_1=2) \xrightarrow{\max} V_{12}\\ \psi_3(x_1=3) \xrightarrow{\max} V_{13}\end{array}\right. \rightarrow \max_x p(x)$$

- backtrack through the argmax values to get $\arg\max_x p(x)$.

noun	noun	noun
verb	verb	verb
adj	adj	adj

ψ_{ij} ψ_1

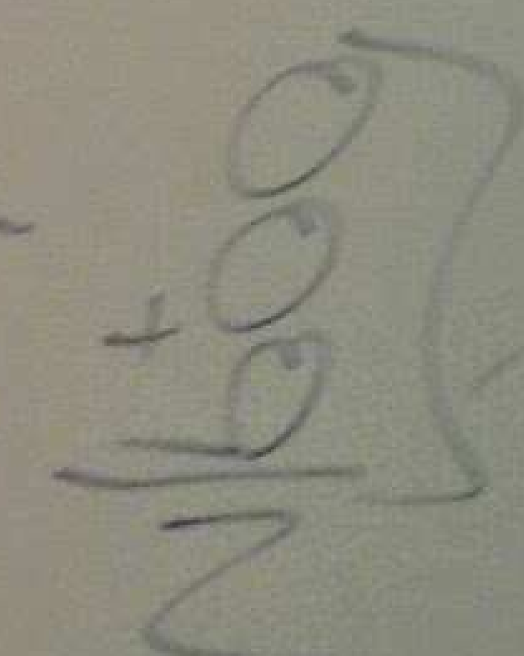
The ball is red.

Forward-Backward

- We can compute Z with a similar algorithm

$$V_{1k} = \psi_1(x_1 = k)$$

$$V_{jk} = \sum_{k'} \psi_{j-1}(x_{j-1} = k', x_j = k)\, V_{(j-1)k'}$$

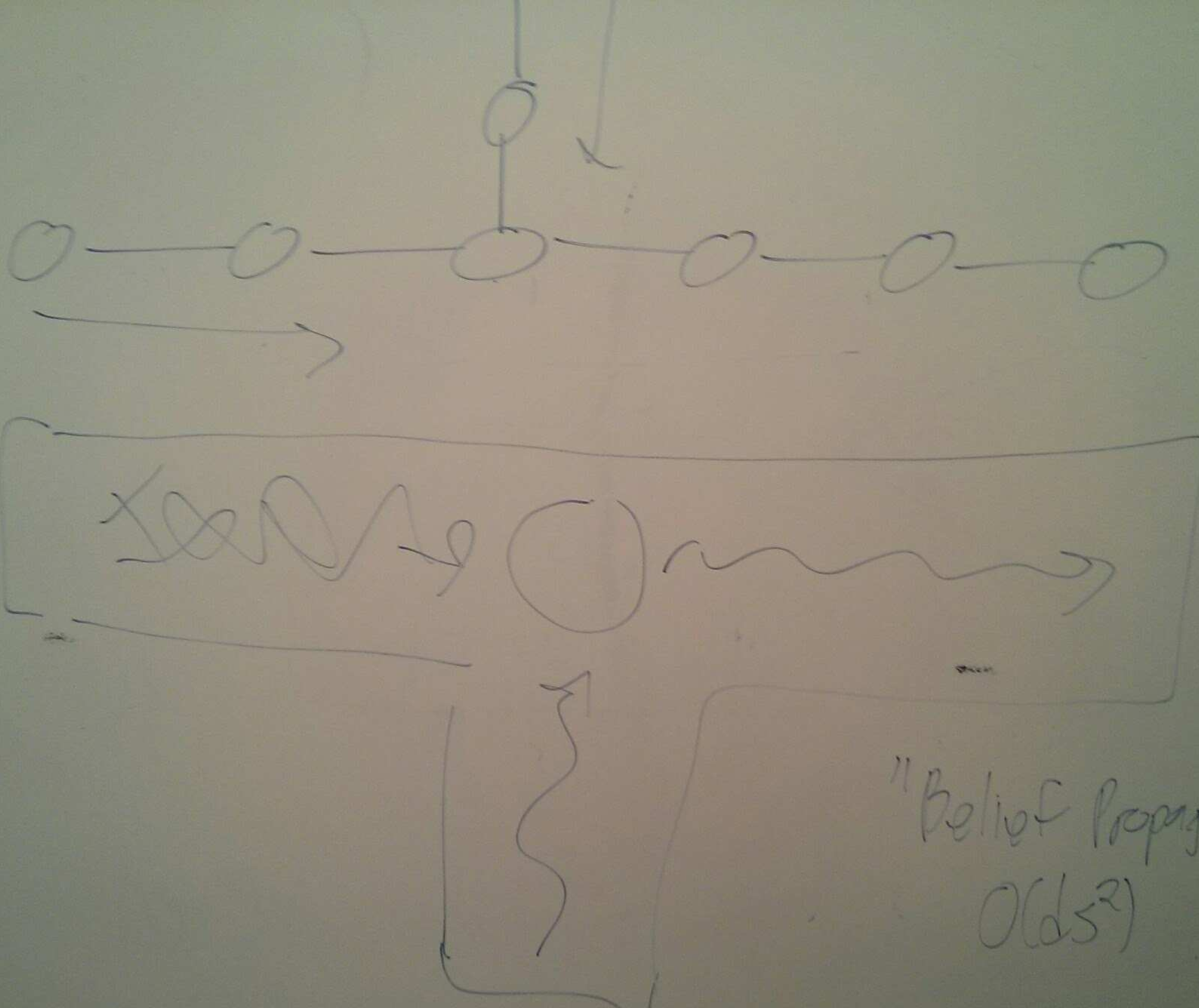


"Backward Algorithm"

$$B_{dk} = 1$$

$$B_{jk} = \sum_{k'} \psi_j(k', k)\, B_{(j+1)k'}$$

$$p(x_j = k) \propto V_{jk} B_{jk}$$



"Belief Propagation"

$O(ds^2)$