

Probabilistic Graphical Models

Lectures 14

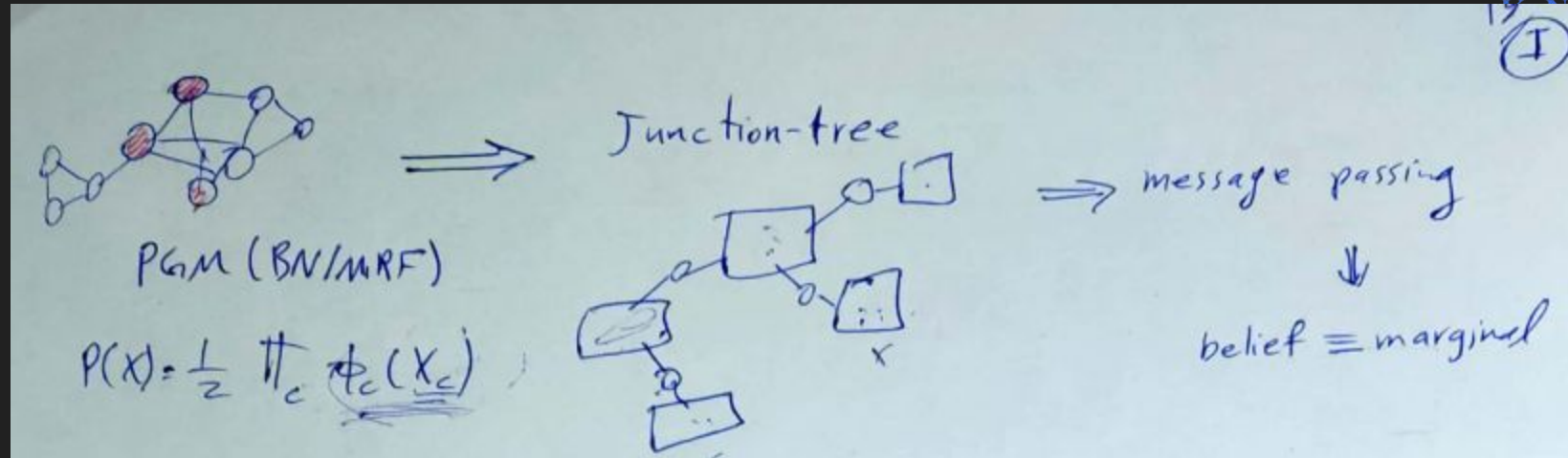
Loopy Belief Propagation

Inference

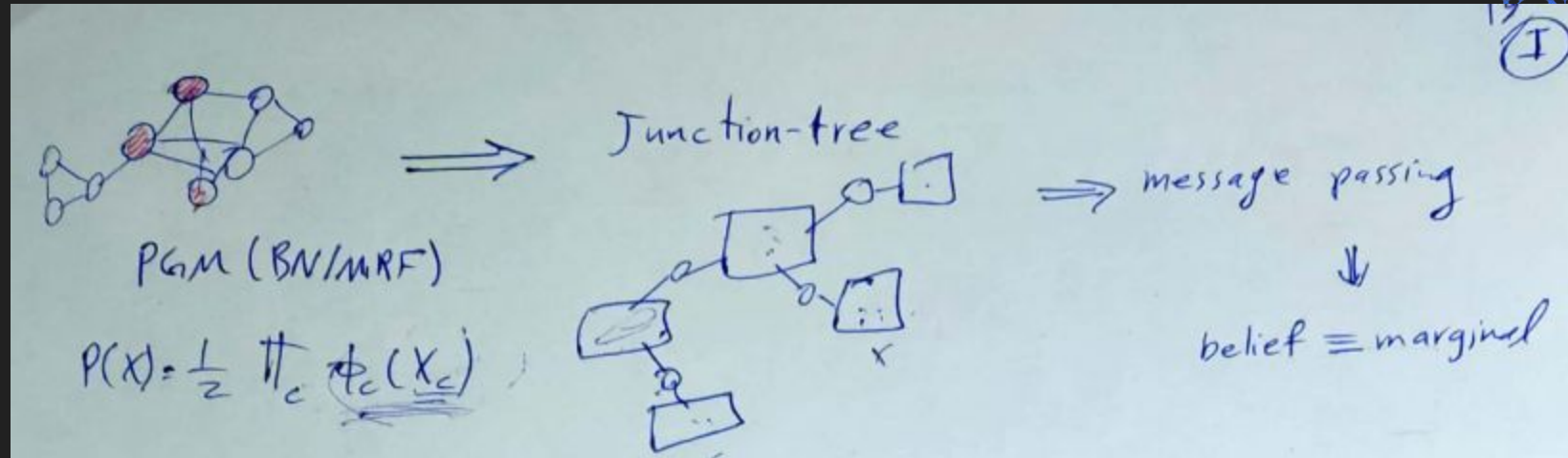


- Exact
 - Variable Elimination
 - Message Passing - Junction Tree
 - Graph-cuts
- Approximate
 - Message Passing - Loopy Belief Propagation
 - Graph-cut based
 - Sampling Based
 - Variational Inference

How to build a cluster tree?



How to build a cluster tree?

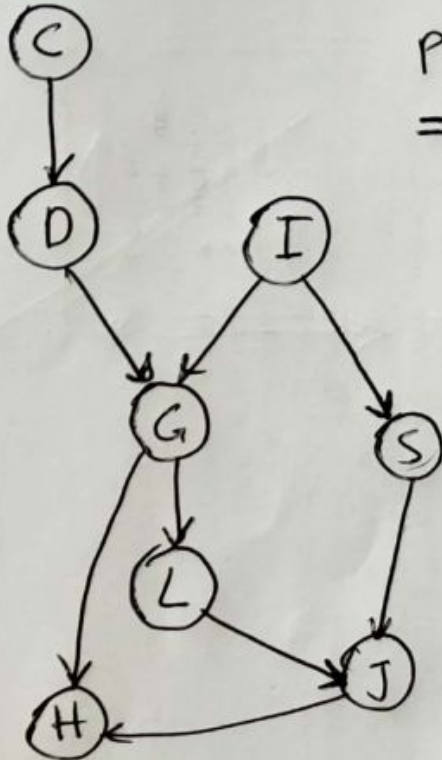


1. Using variable elimination

Build cluster tree using VE



K. N. Toosi
University of Technology



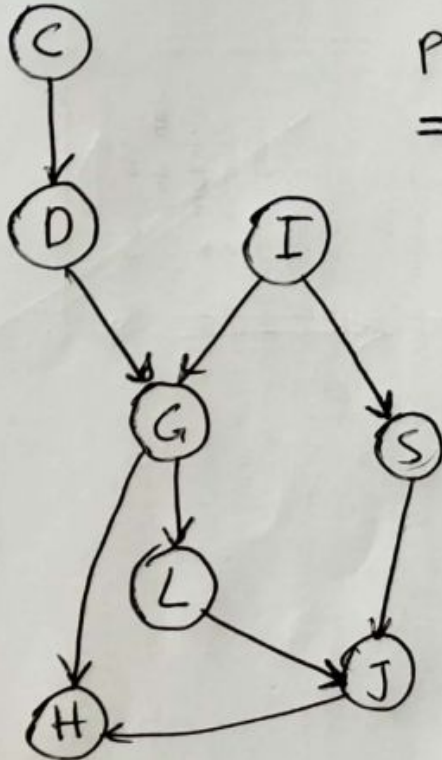
$$\begin{aligned} P(C, D, I, G, L, S, J, H) &= P(C) P(D|C) P(I) P(G|D, I) \\ &\quad P(S|I) P(L|G) P(J|S, L) P(H|J, G) \\ &= \boxed{\phi_1(C)} \boxed{\phi_2(C, D)} \boxed{\phi_3(I)} \boxed{\phi_4(G, D, I)} \\ &\quad \boxed{\phi_5(S, I)} \boxed{\phi_6(L, G)} \boxed{\phi_7(J, S, L)} \\ &\quad \boxed{\phi_8(H, J, G)} \end{aligned}$$

Eliminate: C, D, I, H, G, S, J

Build cluster tree using VE



K. N. Toosi
University of Technology



$$\begin{aligned} P(C, D, I, G, L, S, J, H) &= P(C) P(D|C) P(I) P(G|D, I) \\ &\quad P(S|I) P(L|G) P(J|S, L) P(H|J, G) \\ &= \boxed{\phi_1(C)} \boxed{\phi_2(C, D)} \boxed{\phi_3(I)} \boxed{\phi_4(G, D, I)} \\ &\quad \boxed{\phi_5(S, I)} \boxed{\phi_6(L, G)} \boxed{\phi_7(J, S, L)} \\ &\quad \boxed{\phi_8(H, J, G)} \end{aligned}$$

Eliminate: C, D, I, H, G, S, J

Build cluster tree using VE



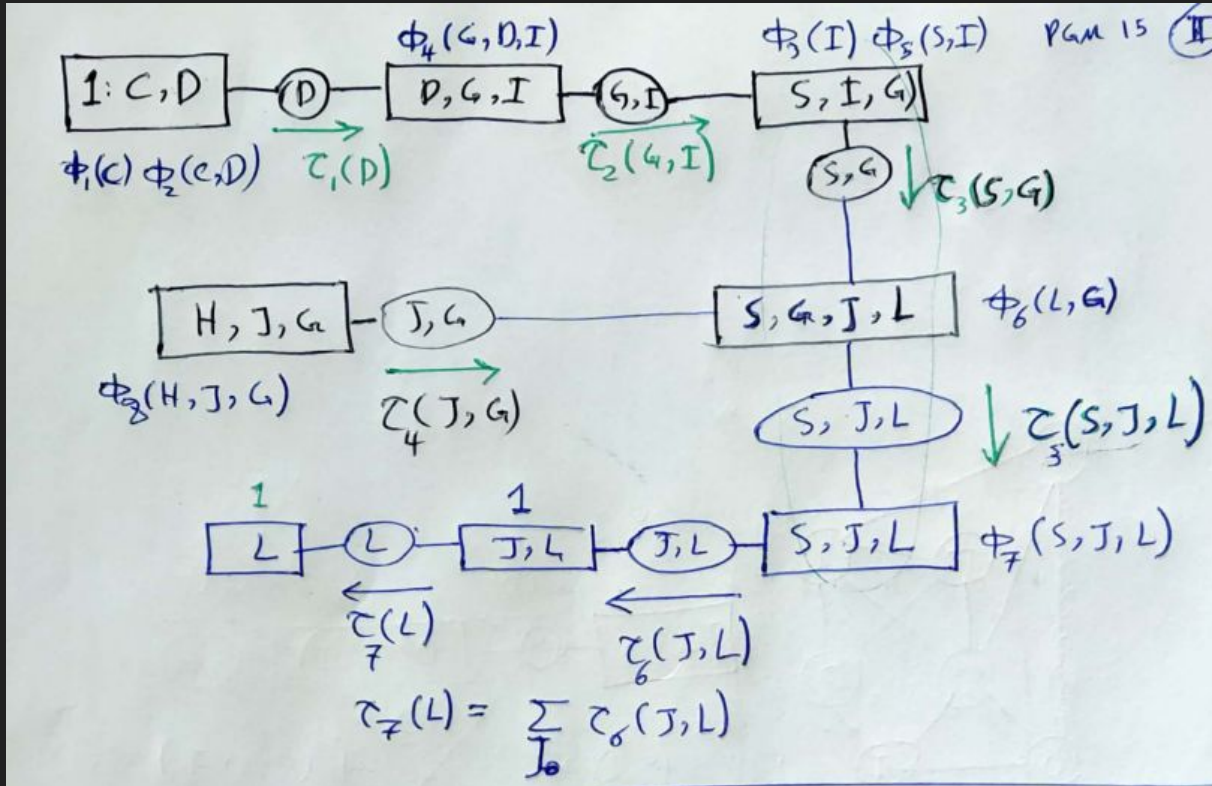
Eliminate C: $\tau_1(D) = \sum_C \phi_1(C) \phi_2(C, D)$

Eliminate D: $\tau_2(G, I) = \sum_D \phi_4(G, D, I) \tau_1(D)$

Eliminate I: $\tau_3(S, G) = \sum_I \phi_3(I) \phi_5(S, I) \tau_2(G, I)$

Eliminate H:

Build cluster tree using VE



How to build a cluster tree?

1. Using variable elimination
2. Determine valid cluster trees



How to build a cluster tree?



Must possess two properties

1- Family Preservation

Each factor can be assigned to some cluster.

for each factor $\phi_c(X_c)$ there must be a cluster C_i with variables X_{C_i} such that $X_c \subseteq X_{C_i}$.

2- Running Intersection Property

for each variable X the subgraph consisting of the cluster-tree containing X ~~is~~ is connected (\equiv also forms a tree).



How to build a cluster tree?

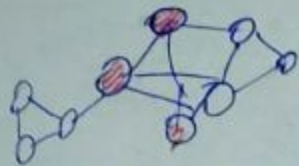
Must possess two properties

1. Family Preservation property: Each factor can be assigned to some cluster.
 - a. For each factor $\varphi_I(X_I)$ there is a cluster C_i such that $X_I \subseteq X_{C_i}$.
2. Running Intersection Property: For each variable X the subgraph containing X is connected (and thus forms a tree).

How to build a cluster tree?

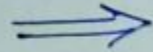


K. N. Toosi
Technology

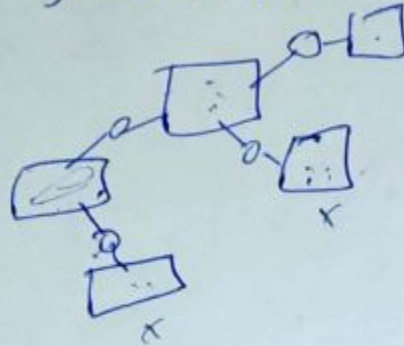


PGM (BN/MRF)

$$P(X) = \frac{1}{Z} \prod_c \phi_c(X_c)$$



Junction-tree



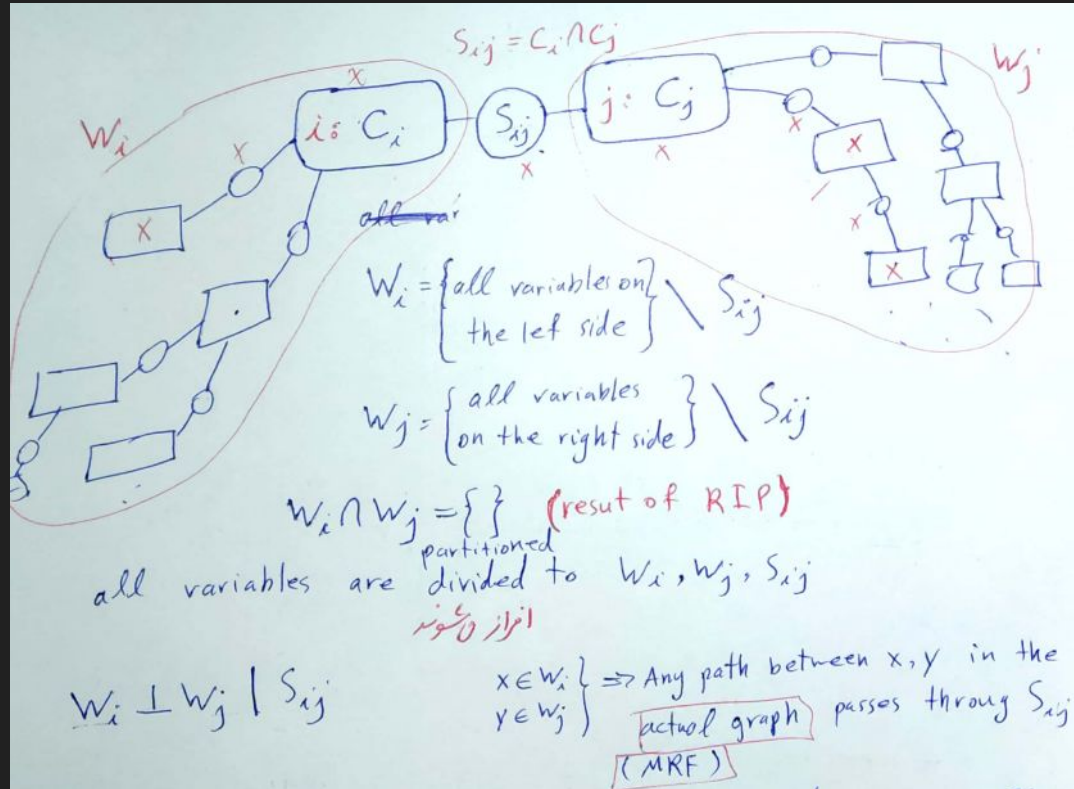
message passing



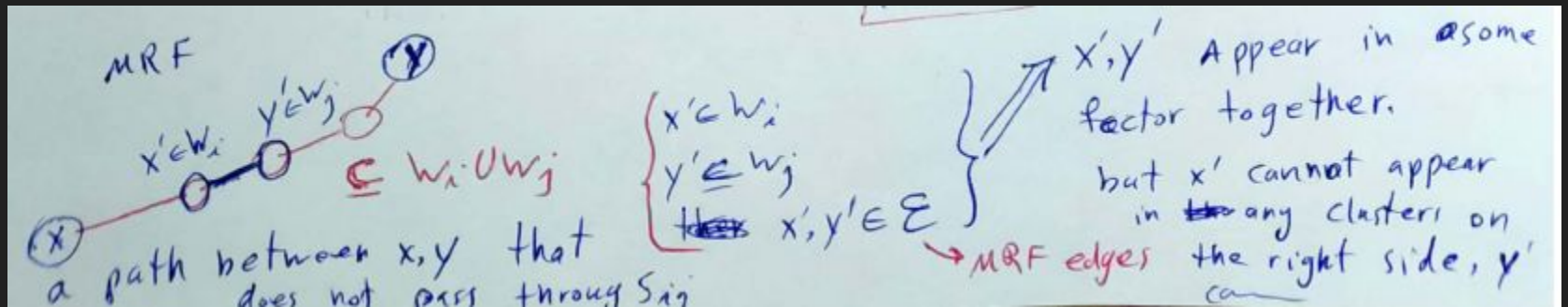
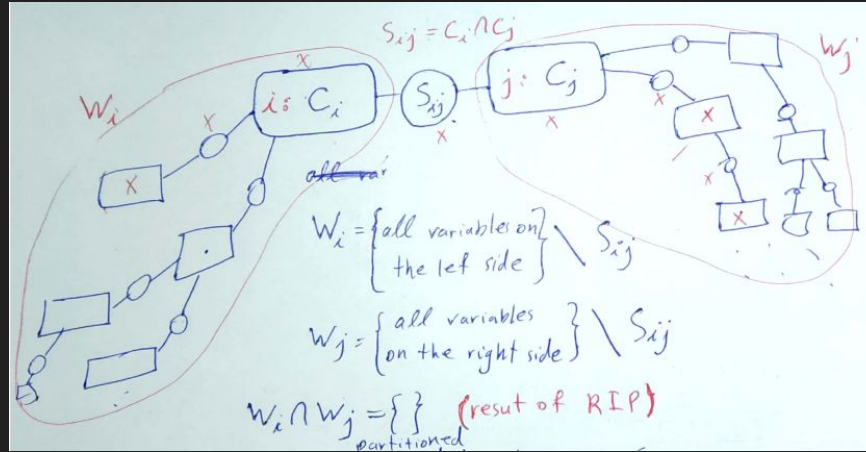
belief \equiv marginal

- 1- Family Preservation Property
- 2- Running Intersection Property (RIP)

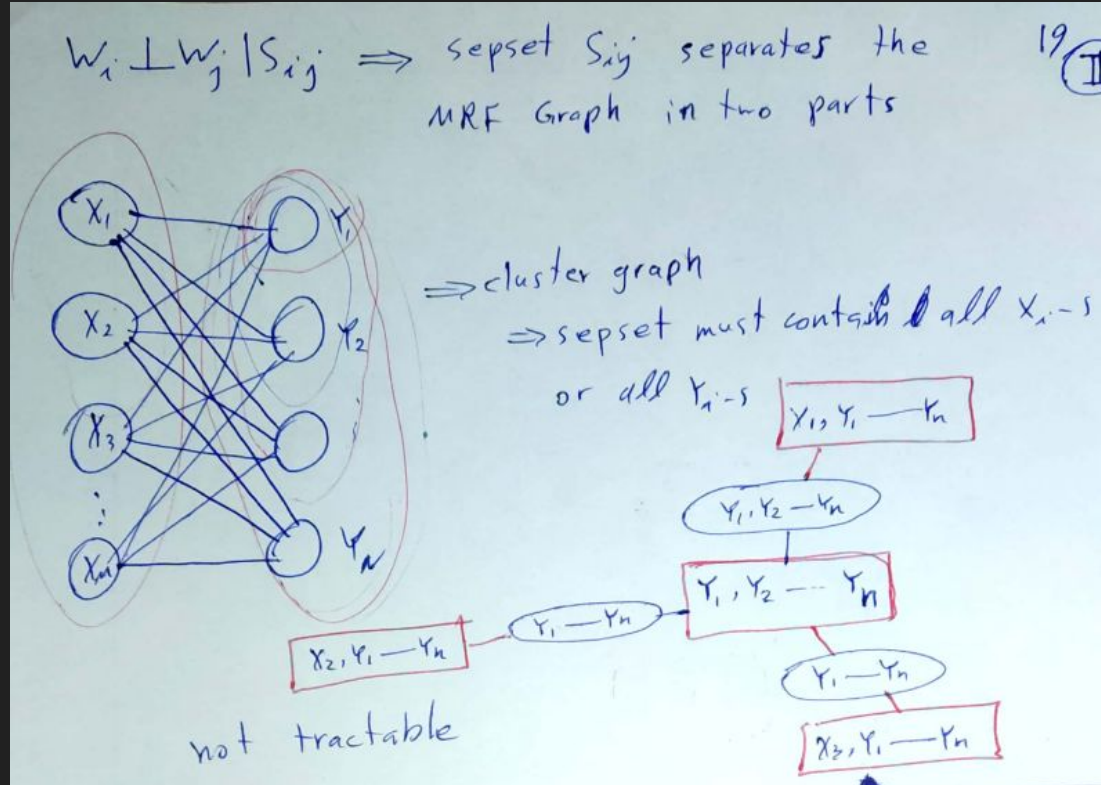
How to build a cluster tree?



How to build a cluster tree?



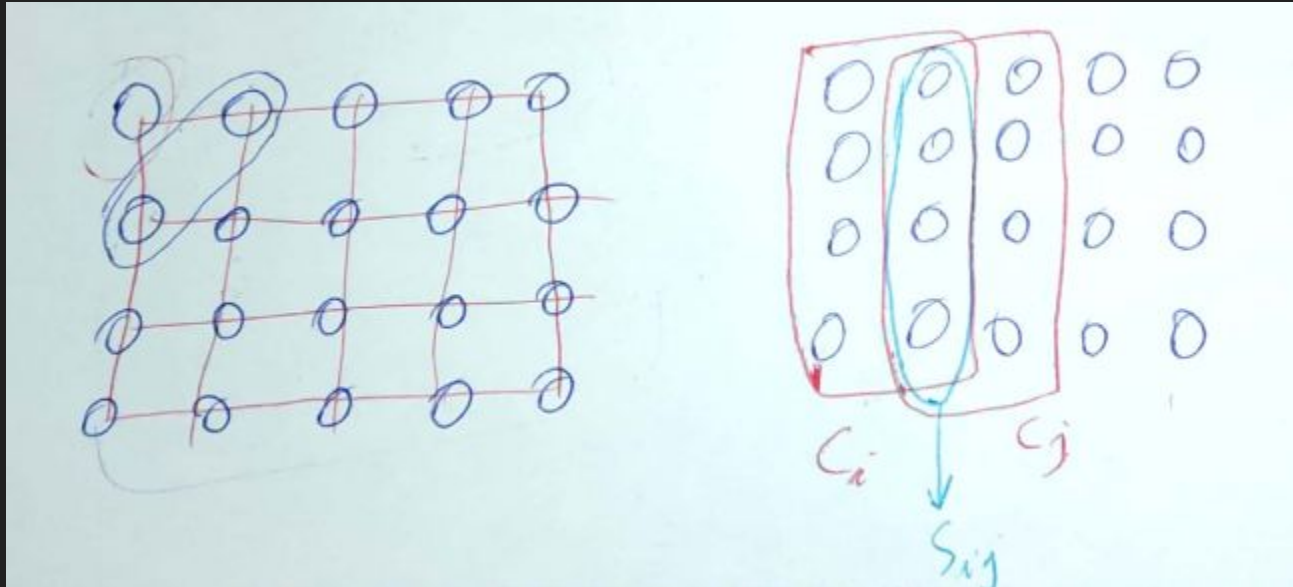
Junction tree - limitations



Junction tree - limitations



K. N. Toosi
University of Technology



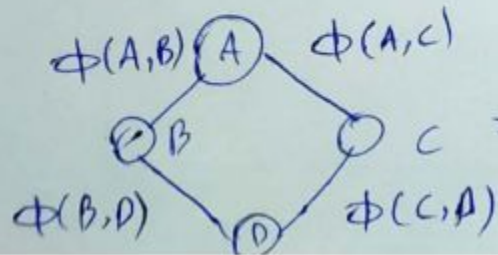
Loopy belief propagation



K. N. Toosi
University of Technology

⇒ Approximate $\left\{ \begin{array}{l} \text{Loopy Belief Propagation} \\ \text{Variational Inference} \\ \text{Sample-Based} \end{array} \right.$

Loopy Belief Propagation: Allow the cluster graph to have loops

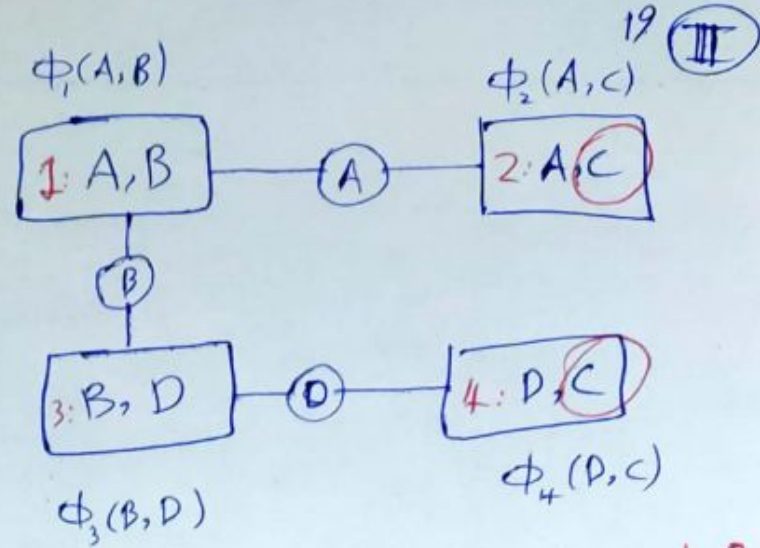
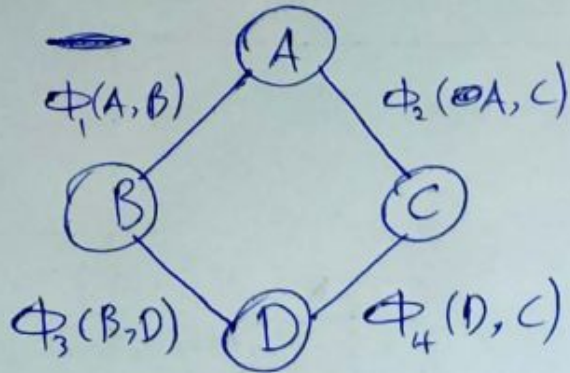


⇒ Can build a junction tree with each cluster having at most 2 ~~variables~~ variables?

Loopy belief propagation



K. N. Toosi
University of Technology



allow loops!

Running intersection Property?

"C" appears in isolated clusters.

Loopy belief propagation



allow loops!

Running intersection Property:
"C" appears in isolated clusters.

$\delta_{3 \rightarrow 1}(B)$

\Rightarrow How to perform message passing? where to start?

Loopy belief propagation

- start from some initial messages, e.g. $\delta_{3 \rightarrow 1}(B) = 1$
- Loop until convergence(?) $\delta(\cdot) = 1$
- pick an edge $i \rightarrow j$ in the cluster graph
- compute (update) the message $\delta_{i \rightarrow j}$

Loopy belief propagation



Loopy belief propagation

- start from some initial messages, e.g. $\delta_{3 \rightarrow \{B\}} = 1$
 $\delta(\cdot) = 1$
- Loop until convergence(?)
- pick an edge $i \rightarrow j$ in the cluster graph
- compute (update) the message $\delta_{i \rightarrow j}$

Messages are computed (updated) more than once

- Does the algorithm converge? *Not Always*
- Does it converge to the right solution? *Usually Not*
- Does it give good approximate solutions? *why?*
Usually gives reasonable solutions

Loopy belief propagation



K. N. Toosi
Technology
IV

Cluster graph Properties ~~for~~

1- Family Preservation (same as before)

2- Running Intersection Property

if x appears in clusters C_i & C_j

\Rightarrow there must be a path ~~where~~ between C_i and C_j
such that x appears in all clusters and
separates along the path

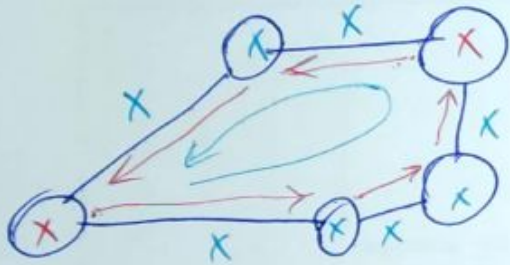
\Rightarrow such a path must be unique.

Loopy belief propagation



K. N. Toosi
University of Technology

for each variable X the subgraph of the cluster graph containing X forms a tree.



~~prevent~~
trying to prevent
falsely reinforcing a belief

In cluster graphs $S_{ij} \subseteq C_i \cap C_j$

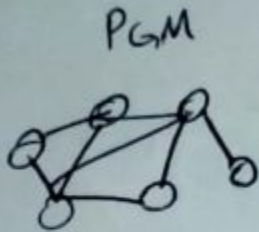
In cluster tree $S_{ij} = C_i \cap C_j$

Loopy belief propagation

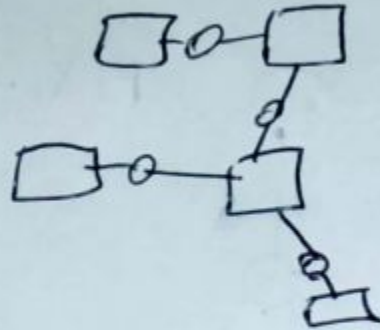


K. N. Toosi
University of Technology

20 (I)



cluster tree



Message passing

⇒ Allow loop in cluster graph ⇒

- iterative message passing
- Approximate Inference

Is there at least one cluster graph?



K. N. Toosi
University of Technology

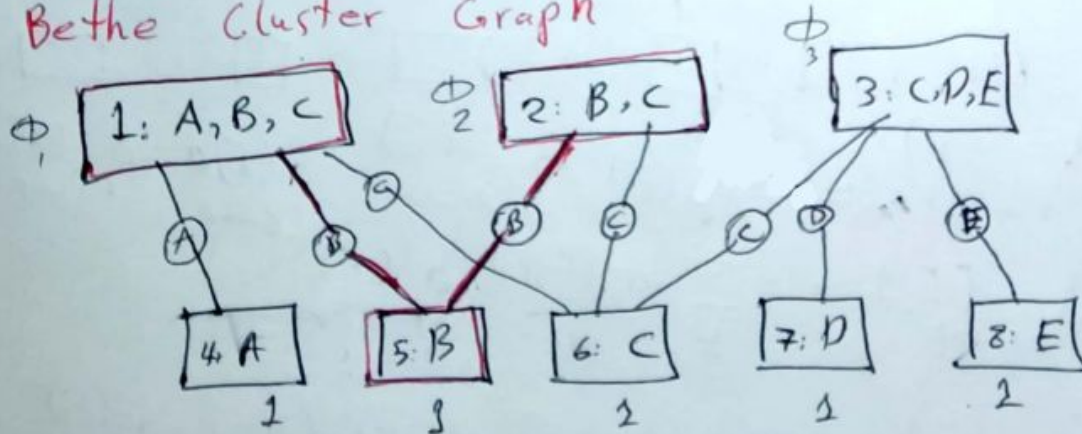
Is there at least one cluster graph?



$$P(A, B, C, D, E) = \phi_1(A, B, C) \phi_2(B, C) \phi_3(C, D, E)$$

is there a cluster graph with FPP & RIP such that the cluster sizes are at most the size of factors?

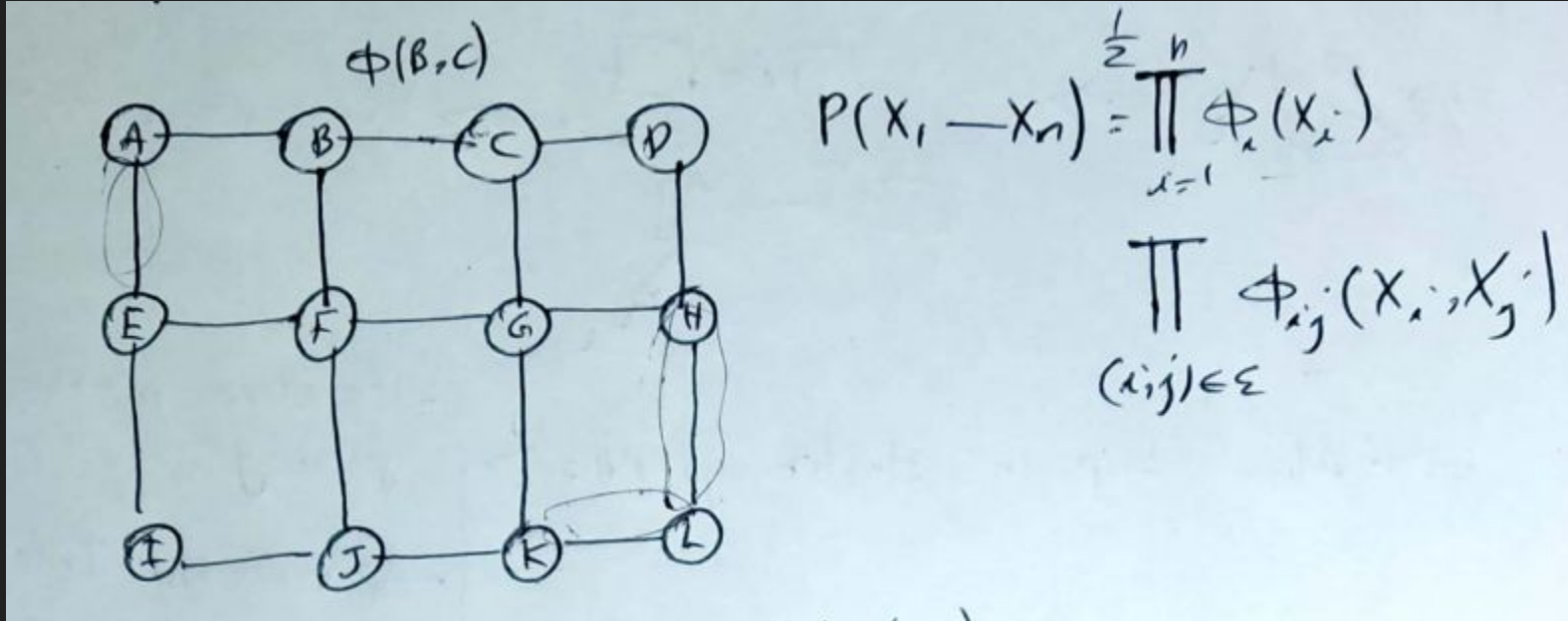
Bethe Cluster Graph



Example: Grids



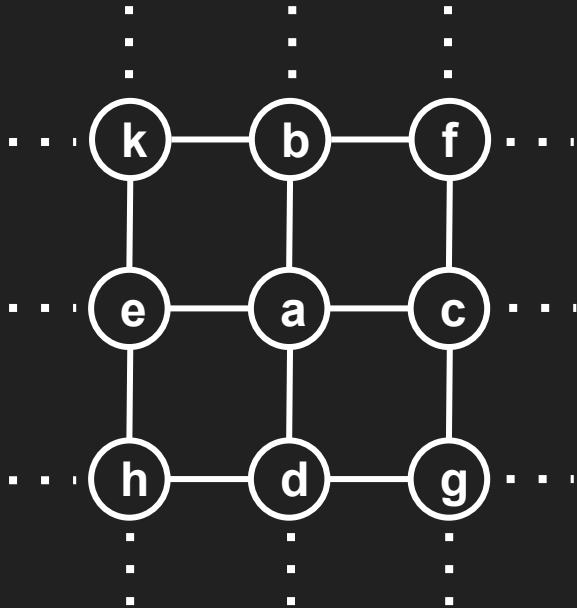
K. N. Toosi
University of Technology



Example: Grids



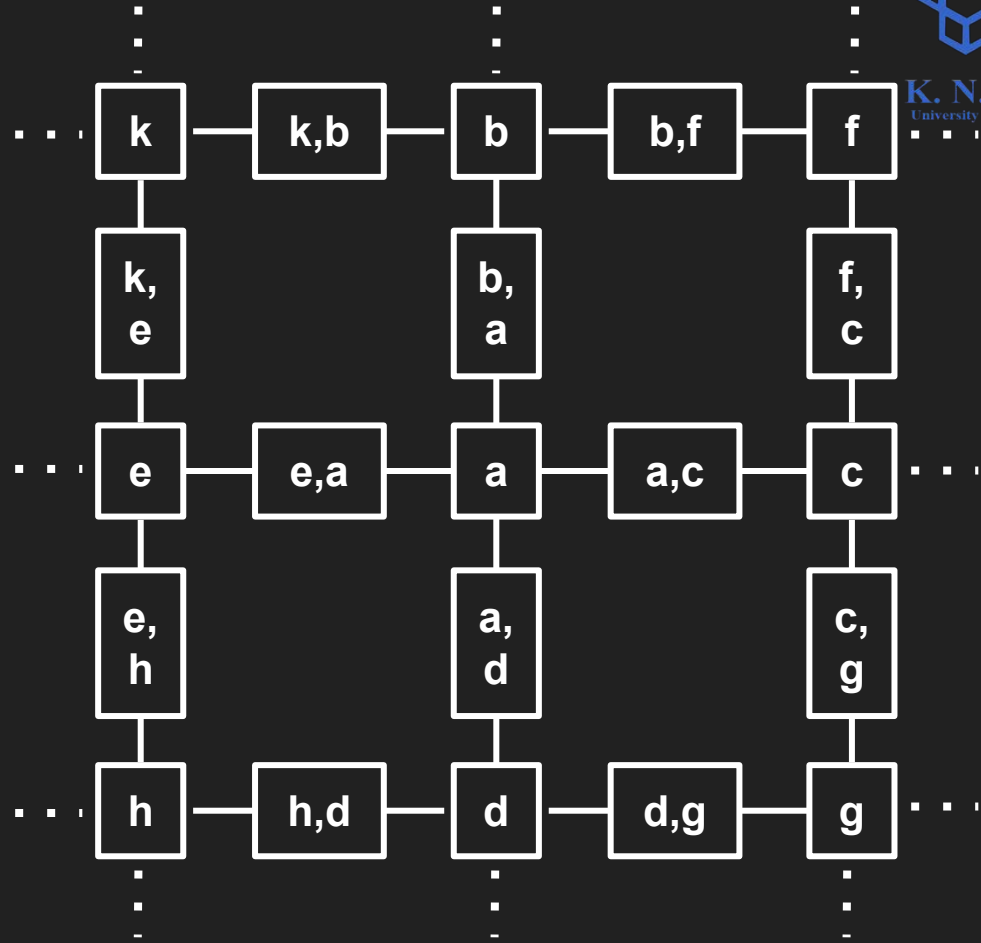
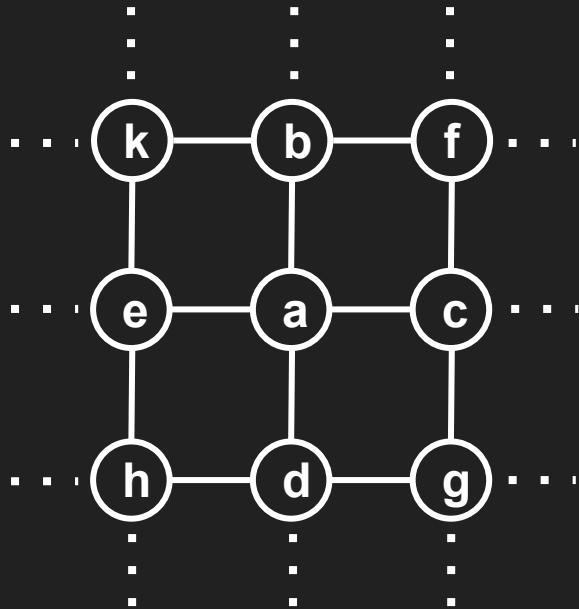
K. N. Toosi
University of Technology



Example: Grids



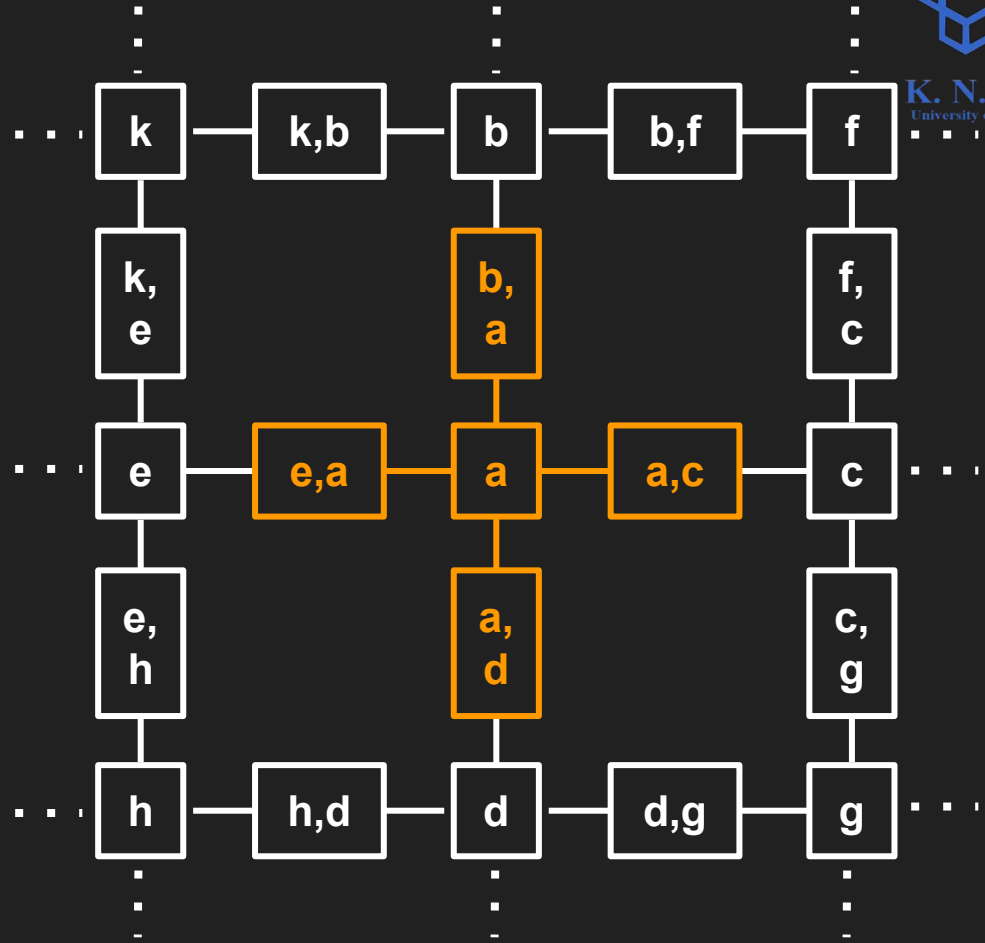
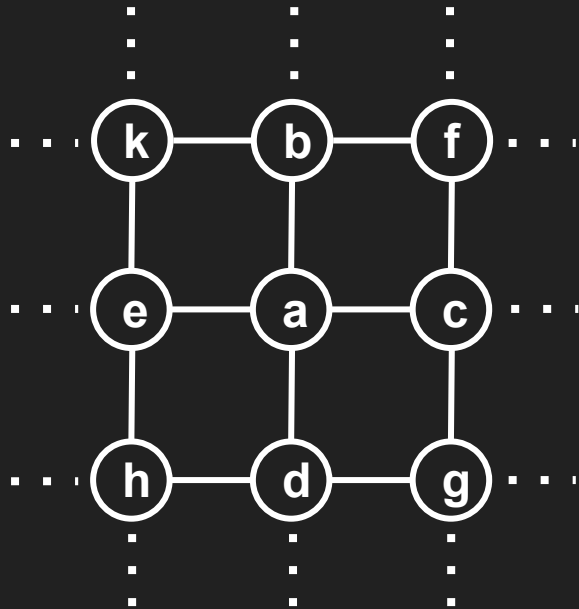
K. N. Toosi
University of Technology



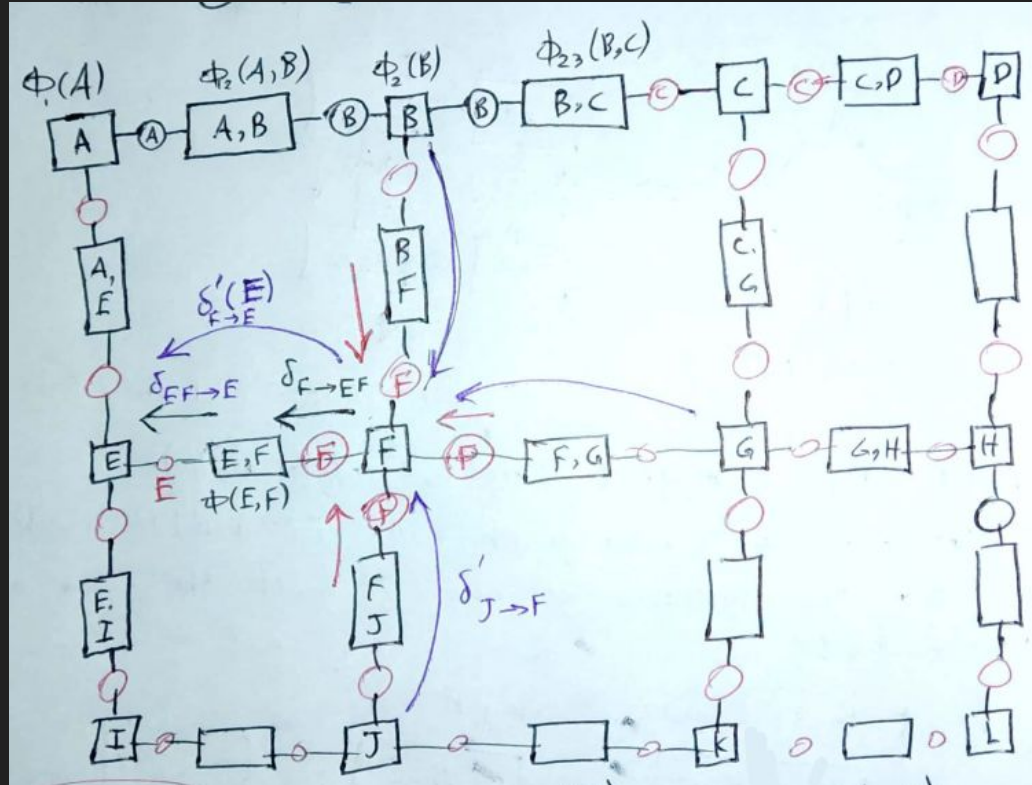
Example: Grids



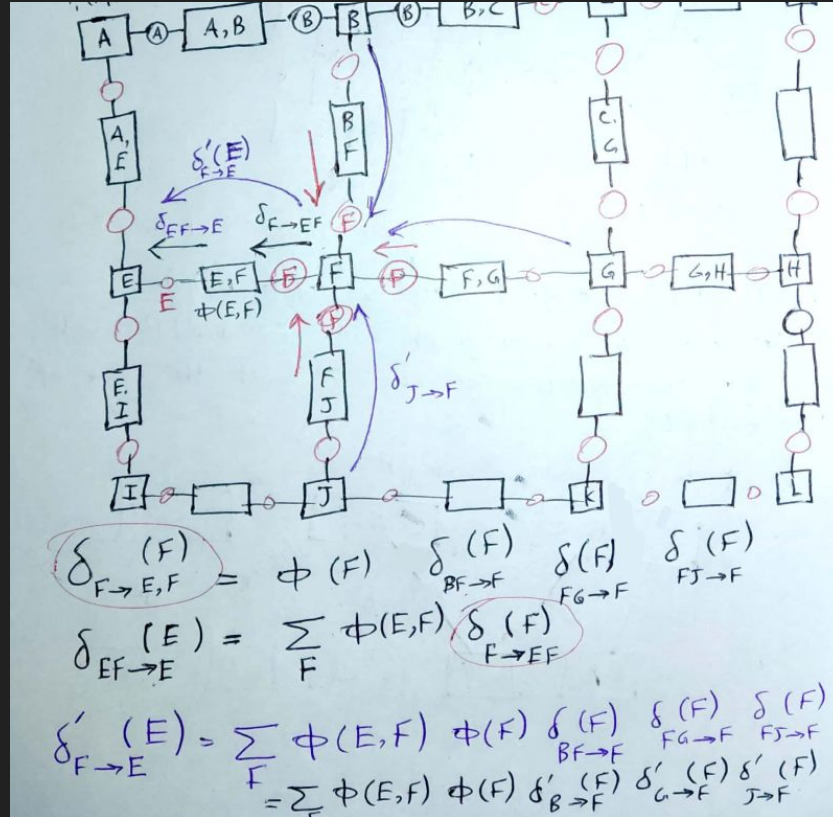
K. N. Toosi
University of Technology



Example: Grids



Example: Grids




Calibration



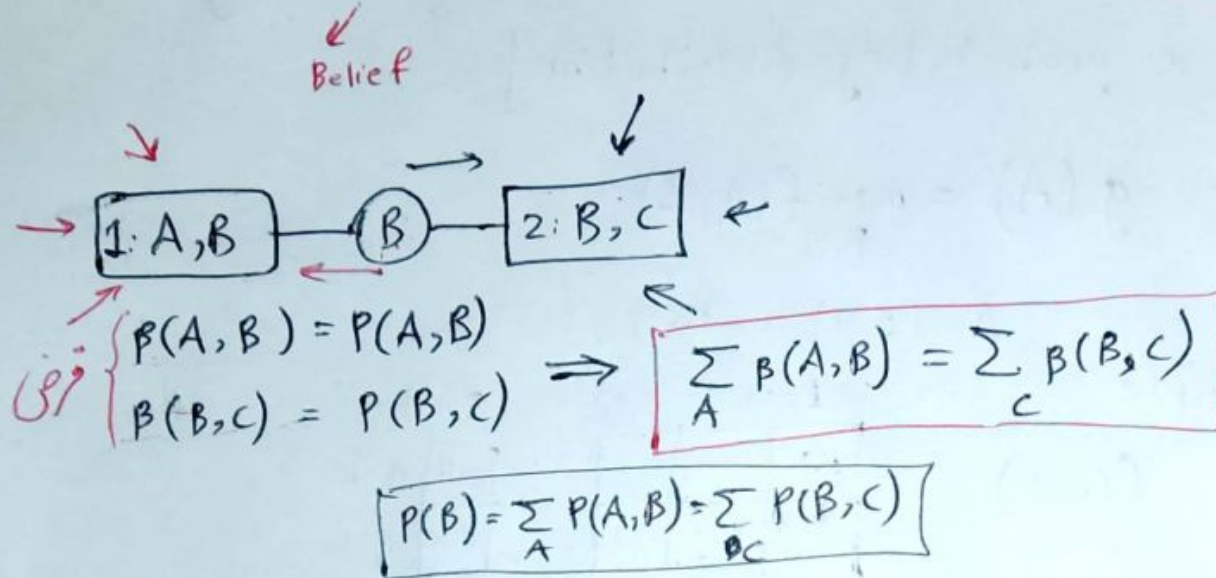
K. N. Toosi
University of Technology

Calibration and Convergence




Assume that message passing gives the exact ²⁰ 

marginals $\beta(A, B) = P(A, B)$

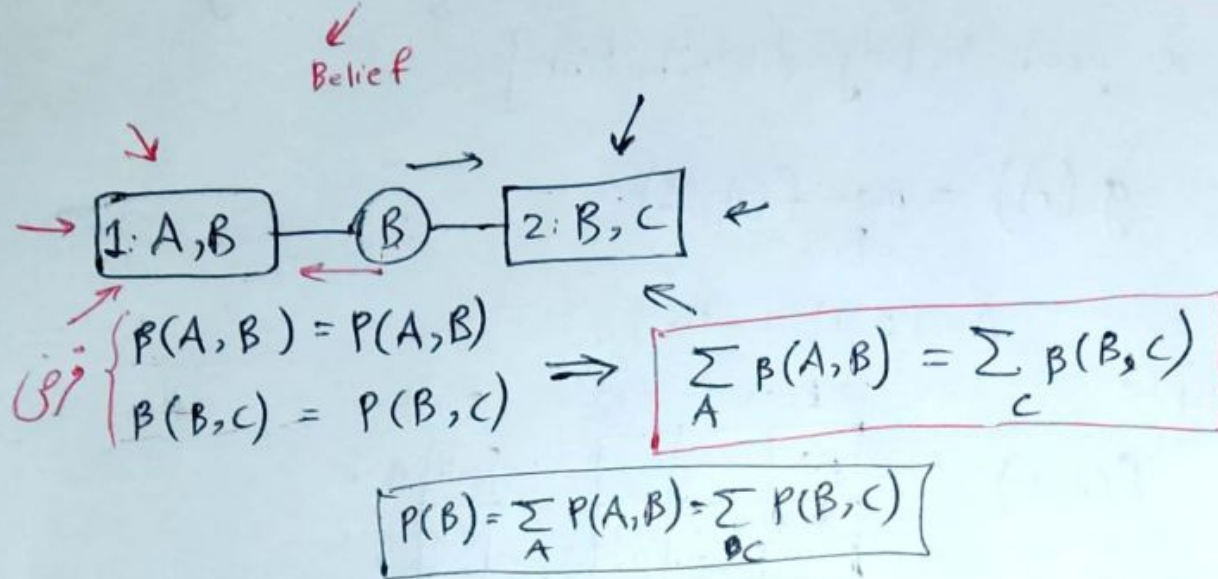


Calibration and Convergence



Assume that message passing gives the exact ²⁰ 

marginals $\beta(A, B) = P(A, B)$



Calibration and Convergence



Diagram illustrating a graphical model with three nodes: Node 1 (A, B), Node B, and Node 2 (B, C). Arrows indicate dependencies: Node 1 depends on A, Node B depends on B, and Node 2 depends on B and C. Bidirectional arrows between nodes 1 and B, and between B and 2, represent message passing.

Handwritten notes below the diagram:

$$\beta(A, B) = P(A, B)$$

$$\beta(B, C) = P(B, C)$$

$$\Rightarrow \sum_A \beta(A, B) = \sum_C \beta(B, C)$$

$$P(B) = \sum_A P(A, B) = \sum_C P(B, C)$$

true solution \Rightarrow beliefs agree upon the setpoint
 (graph is calibrated)

\Leftarrow
 ? NOT
 Necessarily

Graph Calibrated \Rightarrow Message passing has converged.
 Messages do not change anymore.