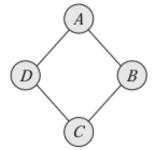
Royal Institute of Technology

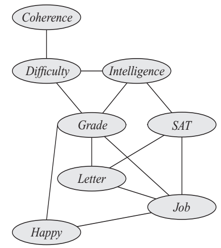
STAT. METH. IN CS –EXACT INFERENCE, CH 20

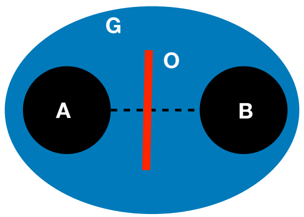
Lecture 8

UGM

- UGMs - Undirected graphical models
- What is the direction between 2 pixels, 2 proteins?
- Probabilistic interpretation?
- p factorizes over G – can be expressed as normalized product over factors associated with cliques
- here categorical







SEPARATION AND CI OF UGM

★ A is separated from B given O in G if there is no path between A and B in G\O

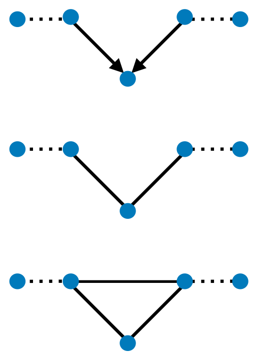
★ In a graph G,

↕

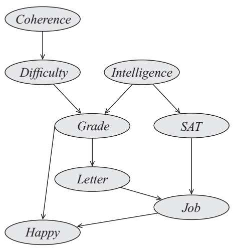
A is separated from B given O

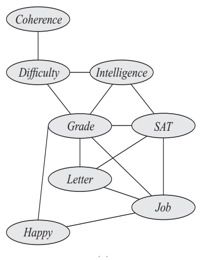
DEF I-MAP

- G is an I-map for p if all independence relation in G hold for p, i.e., I(G)⊆I(p)
- Moralize add edge between any two parents
- We can moralize a DGM and get a UGM having no more independence relations
- Each family has a cliques in the moralized UGM



CONVERTING DGM TO UGM & SUING JUNCTION TREE





- Moralize and remove directions
  - does not introduce new independencies!
- Use CPDs as factors

- Families are cliques
- Cliques in bags of any Junction tree
- We will work with marginals of bags

A FACTOR PRODUCT

a^1	b^1	0.5
a^1	b^2	0.8
a^2	b^1	0.1
a^2	b^2	0
a^3	b^1	0.3
a^3	b^2	0.9

b^1	c^1	0.5
b^1	c^2	0.7
b^2	c^1	0.1
b^2	c^2	0.2

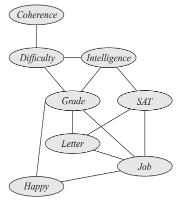
a^1	b^1	c^1	$0.5\cdot0.5 = 0.25$
a^1	b^1	c^2	$0.5\cdot0.7 = 0.35$
a^1	b^2	c^1	$0.8\cdot0.1 = 0.08$
a^1	b^2	c^2	$0.8\cdot0.2 = 0.16$
a^2	b^1	c^1	$0.1\cdot0.5 = 0.05$
a^2	b^1	c^2	$0.1\cdot0.7 = 0.07$
a^2	b^2	c^1	$0\cdot0.1 = 0$
a^2	b^2	c^2	$0\cdot0.2 = 0$
a^3	b^1	c^1	$0.3\cdot0.5 = 0.15$
a^3	b^1	c^2	$0.3\cdot0.7 = 0.21$
a^3	b^2	c^1	$0.9\cdot0.1 = 0.09$
a^3	b^2	c^2	$0.9\cdot0.2 = 0.18$

SUMMING OUT - IN FACTOR

a^1	b^1	c^1	0.25
a^1	b^1	c^2	0.35
a^1	b^2	c^1	0.08
a^1	b^2	c^2	0.16
a^2	b^1	c^1	0.05
a^2	b^1	c^2	0.07
a^2	b^2	c^1	0
a^2	b^2	c^2	0
a^3	b^1	c^1	0.15
a^3	b^1	c^2	0.21
a^3	b^2	c^1	0.09
a^3	b^2	c^2	0.18

a^1	c^1	0.33
a^1	c^2	0.51
a^2	c^1	0.05
a^2	c^2	0.07
a^3	c^1	0.24
a^3	c^2	0.39

JOINT AS FACTOR PRODUCT



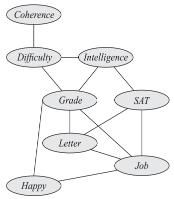
-

$$P(C,D,I,G,S,L,J,H) = P(C)P(D|C)P(I)P(G|I,D)P(S|I)P(L|G)P(J|L,S)P(H|G,J)$$

-

$$p(C,D,I,G,S,L,J,H) = \psi_C(C)\psi_D(D,C)\psi_I(I)\psi_G(G,I,D)\psi_S(S,I)\psi_L(L,G)\psi_J(J,L,S)\psi_H(H,G,J)$$

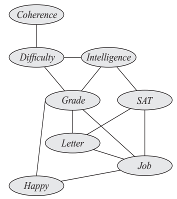
SUMMING OUT



•

$$p(J) = \sum_L \sum_S \sum_G \sum_H \sum_I \sum_D \sum_C p(C, D, I, G, S, L, J, H)$$

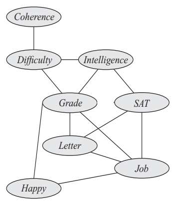
SUMMING OUT



•

$$\begin{aligned} p(J) &= \sum_{L,S,G,H,I,D,C} p(C, D, I, G, S, L, J, H) \\ &= \sum_{L,S,G,H,I,D,C} \psi_C(C)\psi_D(D,C)\psi_I(I)\psi_G(G,I,D)\psi_S(S,I)\psi_L(L,G) \\ &\quad \times \psi_J(J,L,S)\psi_H(H,G,J) \\ &= \sum_{L,S} \psi_J(J,L,S) \sum_G \psi_L(L,G) \sum_H \psi_H(H,G,J) \sum_I \psi_S(S,I)\psi_I(I) \\ &\quad \times \sum_D \psi_G(G,I,D) \sum_C \psi_C(C)\psi_D(D,C) \end{aligned}$$

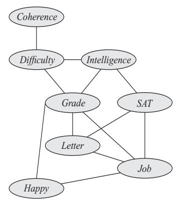
SUMMING OUT



•

$$\begin{aligned} p(J) &= \sum_{L,S} \psi_J(J,L,S) \sum_G \psi_L(L,G) \sum_H \psi_H(H,G,J) \sum_I \psi_S(S,I)\psi_I(I) \\ &\quad \times \sum_D \psi_G(G,I,D) \sum_C \psi_C(C)\psi_D(D,C) \end{aligned}$$

SUMMING OUT



$$\begin{aligned} p(J) &= \sum_{L,S} \psi_J(J,L,S) \sum_G \psi_L(L,G) \sum_H \psi_H(H,G,J) \sum_I \psi_S(S,I)\psi_I(I) \\ &\quad \times \sum_D \psi_G(G,I,D) \sum_C \psi_C(C)\psi_D(D,C) \end{aligned}$$

Temporary factors

$$\tau_1'(C, D) = \psi_C(C)\psi_D(D,C)$$

$$\tau_1(D) = \sum_C \tau_1'(C, D)$$

THE ENTIRE SEQUENCE

$$\sum_L\sum_S \psi_J(J,L,S)\sum_G\psi_L(L,G)\sum_H\psi_H(H,G,J)\sum_I\psi_S(S,I)\psi_I(I)\sum_D\psi_G(G,I,D)\underbrace{\sum_C\psi_C(C)\psi_D(D,C)}_{\tau_1(D)}$$

$$\sum_L\sum_S \psi_J(J,L,S)\sum_G\psi_L(L,G)\sum_H\psi_H(H,G,J)\sum_I\psi_S(S,I)\psi_I(I)\underbrace{\sum_D\psi_G(G,I,D)\tau_1(D)}_{\tau_2(G,I)}$$

$$\sum_L\sum_S \psi_J(J,L,S)\sum_G\psi_L(L,G)\sum_H\psi_H(H,G,J)\underbrace{\sum_I\psi_S(S,I)\psi_I(I)\tau_2(G,I)}_{\tau_3(G,S)}$$

$$\sum_L\sum_S \psi_J(J,L,S)\sum_G\psi_L(L,G)\underbrace{\sum_H\psi_H(H,G,J)}_{\tau_4(G,J)}\tau_3(G,S)$$

$$\sum_L\sum_S \psi_J(J,L,S)\underbrace{\sum_G\psi_L(L,G)\tau_4(G,J)\tau_3(G,S)}_{\tau_5(J,L,S)}$$

$$\sum_L\underbrace{\sum_S \psi_J(J,L,S)\tau_5(J,L,S)}_{\tau_6(J,L)}$$

$$\underbrace{\sum_L\tau_6(J,L)}_{\tau_7(J)}$$

A FILL-IN EDGE

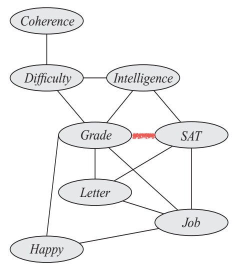
$$\sum_L\sum_S \psi_J(J,L,S)\sum_G\psi_L(L,G)\sum_H\psi_H(H,G,J)\underbrace{\sum_I\psi_S(S,I)\psi_I(I)\tau_2(G,I)}_{\tau_3(G,S)}$$

$$\sum_L\sum_S \psi_J(J,L,S)\sum_G\psi_L(L,G)\underbrace{\sum_H\psi_H(H,G,J)}_{\tau_4(G,J)}\tau_3(G,S)$$

$$\sum_L\sum_S \psi_J(J,L,S)\underbrace{\sum_G\psi_L(L,G)\tau_4(G,J)\tau_3(G,S)}_{\tau_5(J,L,S)}$$

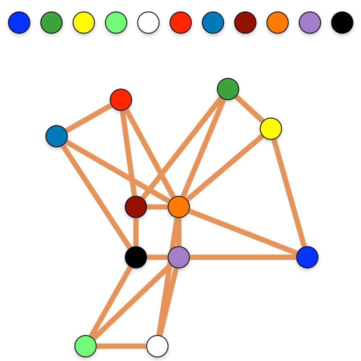
$$\sum_L\underbrace{\sum_S \psi_J(J,L,S)\tau_5(J,L,S)}_{\tau_6(J,L)}$$

$$\underbrace{\sum_L\tau_6(J,L)}_{\tau_7(J)}$$



Fill in edge between two vertices in a "temporary factor"

ALTERNATIVE DEFINITION OF WIDTH



Elimination order of G – ordering of V(G)

$v_1,\dots,v_n$

It defines a sequence of graphs $G=G_0,\dots,G_n$

where $G_i := G_{i-1}\setminus v_i \cup \{(u,w): u,w\in N_{G_{i-1}}(v_i)\}$

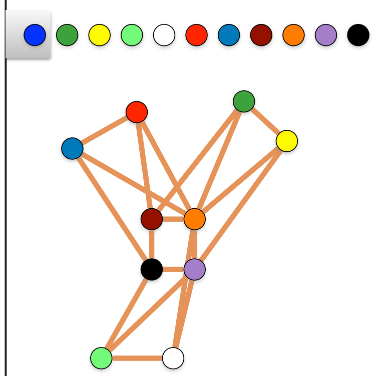
(i.e., we add edges between all neighbours)

width of order = (max clique size in $\cup_i G_i$) -1

= $\max_i$ degree of v_i in G_i

width of G = min width of order

ALTERNATIVE DEFINITION OF WIDTH



Elimination order of G – ordering of V(G)

$v_1,\dots,v_n$

It defines a sequence of graphs $G=G_1,\dots,G_n$

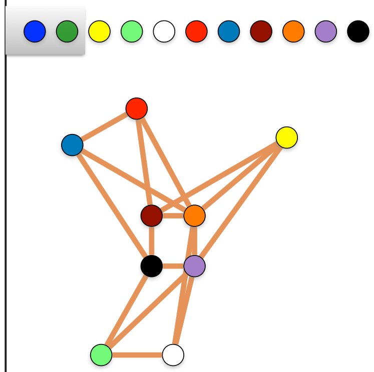
where $G_i := G_{i-1}\setminus v_i \cup \{(u,w): u,w\in N_{G_{i-1}}(v_i)\}$

(i.e., we add edges between all neighbours)

width of order = (max clique size in $\cup_i G_i$) -1

= $\max_i$ degree of v_i in G_i

width of G = min width of order



ALTERNATIVE DEFINITION OF WIDTH

Elimination order of G – ordering of V(G)

$v_1,\ldots,v_n$

It defines a sequence of graphs $G=G_1,\ldots,G_n$

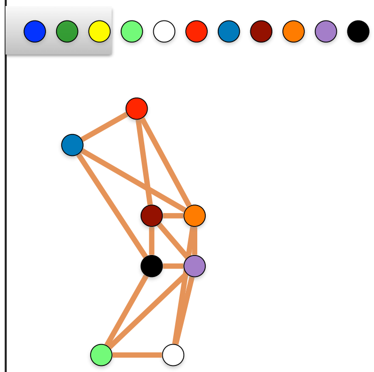
where $G_i := G_{i-1} \setminus v_i \cup \{(u,w): u,w \in N_{G_{i-1}}(v_i)\}$

(i.e., we add edges between all neighbours)

width of order $= (\text{max clique size in } \cup_i G_i) - 1$

$= \max_i \text{degree of } v_i \text{ in } G_i$

width of G = min width of order



ALTERNATIVE DEFINITION OF WIDTH

Elimination order of G – ordering of V(G)

$v_1,\ldots,v_n$

It defines a sequence of graphs $G=G_1,\ldots,G_n$

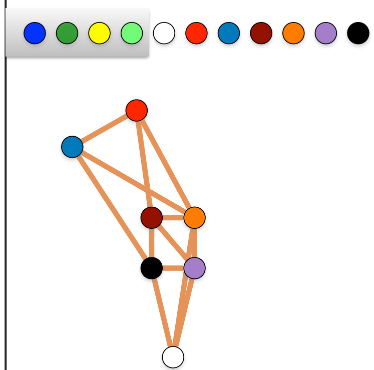
where $G_i := G_{i-1} \setminus v_i \cup \{(u,w): u,w \in N_{G_{i-1}}(v_i)\}$

(i.e., we add edges between all neighbours)

width of order $= (\text{max clique size in } \cup_i G_i) - 1$

$= \max_i \text{degree of } v_i \text{ in } G_i$

width of G = min width of order



ALTERNATIVE DEFINITION OF WIDTH

Elimination order of G – ordering of V(G)

$v_1,\ldots,v_n$

It defines a sequence of graphs $G=G_1,\ldots,G_n$

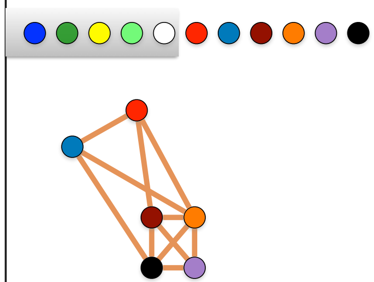
where $G_i := G_{i-1} \setminus v_i \cup \{(u,w): u,w \in N_{G_{i-1}}(v_i)\}$

(i.e., we add edges between all neighbours)

width of order $= (\text{max clique size in } \cup_i G_i) - 1$

$= \max_i \text{degree of } v_i \text{ in } G_i$

width of G = min width of order



ALTERNATIVE DEFINITION OF WIDTH

Elimination order of G – ordering of V(G)

$v_1,\ldots,v_n$

It defines a sequence of graphs $G=G_1,\ldots,G_n$

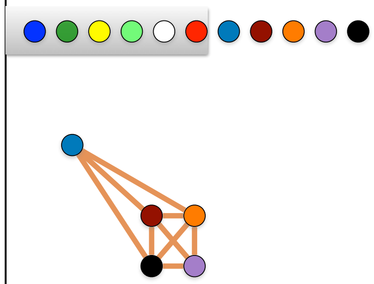
where $G_i := G_{i-1} \setminus v_i \cup \{(u,w): u,w \in N_{G_{i-1}}(v_i)\}$

(i.e., we add edges between all neighbours)

width of order $= (\text{max clique size in } \cup_i G_i) - 1$

$= \max_i \text{degree of } v_i \text{ in } G_i$

width of G = min width of order



ALTERNATIVE DEFINITION OF WIDTH

Elimination order of G – ordering of V(G)

$v_1,\ldots,v_n$

It defines a sequence of graphs $G=G_1,\ldots,G_n$

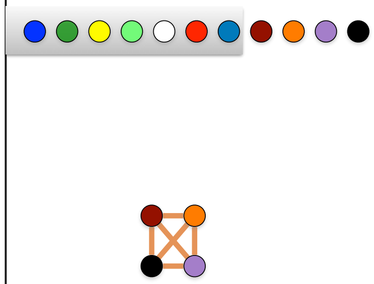
where $G_i := G_{i-1} \setminus v_i \cup \{(u,w): u,w \in N_{G_{i-1}}(v_i)\}$

(i.e., we add edges between all neighbours)

width of order $=$ (max clique size in $\cup_i G_i$) -1

$=$ $\max_i$ degree of v_i in G_i

width of G = min width of order



ALTERNATIVE DEFINITION OF WIDTH

Elimination order of G – ordering of V(G)

$v_1,\ldots,v_n$

It defines a sequence of graphs $G=G_1,\ldots,G_n$

where $G_i := G_{i-1} \setminus v_i \cup \{(u,w): u,w \in N_{G_{i-1}}(v_i)\}$

(i.e., we add edges between all neighbours)

width of order $=$ (max clique size in $\cup_i G_i$) -1

$=$ $\max_i$ degree of v_i in G_i

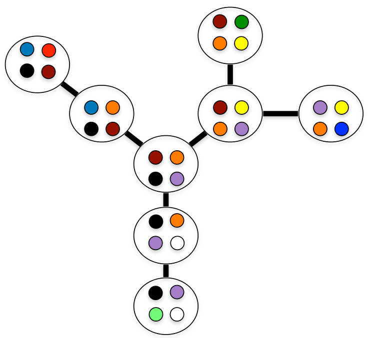
width of G = min width of order

JUNCTION TREE

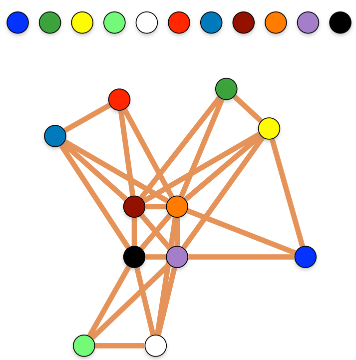
Definition junction tree

- Each vertex of G in some bag
- Each edge of G in some bag
- Each vertex of G induce a subtree (running intersection property)
- Width is (Size of largest bag)-1

T,B junction tree

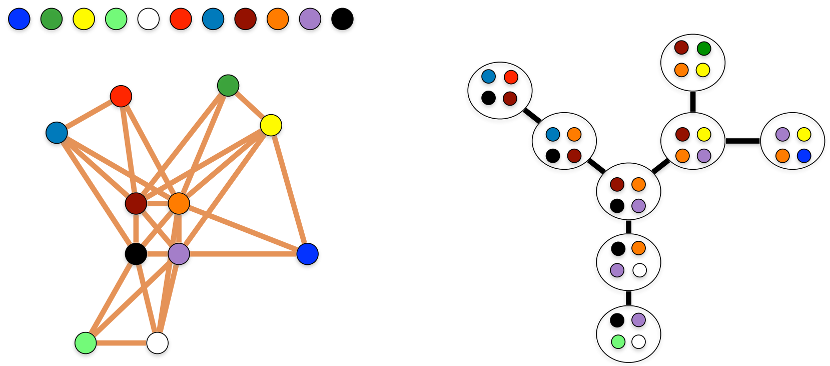


ALTERNATIVE DEFINITION OF WIDTH



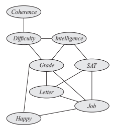
A graph has a junction tree of width k if and only if it has a width k elimination order

(i.e., eliminated vertices have at most k neighbours when eliminated).



JUNCTION TREE AND ELIMINATION ORDER

COMPLEXITY OF VERTEX ELIMINATION



$$\sum_L \sum_S \psi_J(J,L,S) \sum_G \psi_L(L,G) \sum_H \psi_H(H,G,J) \underbrace{\sum_I \psi_S(S,I)\psi_I(I)\tau_2(G,I)}_{\tau_3(G,S)}$$

$$\sum_L \sum_S \psi_J(J,L,S) \sum_G \psi_L(L,G) \underbrace{\sum_H \psi_H(H,G,J)}_{\tau_4(G,J)} \tau_3(G,S)$$

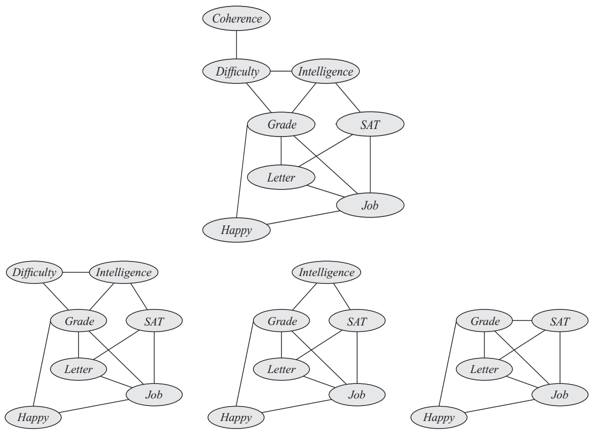
$$\sum_L \sum_S \psi_J(J,L,S) \underbrace{\sum_G \psi_L(L,G)\tau_4(G,J)\tau_3(G,S)}_{\tau_5(J,L,S)}$$

$$\sum_L \underbrace{\sum_S \psi_J(J,L,S)\tau_5(J,L,S)}_{\tau_6(J,L)}$$

$$\underbrace{\sum_L \tau_6(J,L)}_{\tau_7(J)}$$

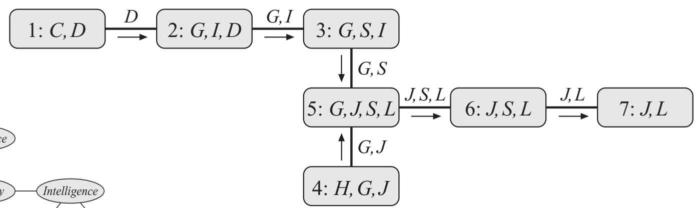
- If the width is w
- Time O(|V(G)| 2^w) for binary
- For categorical with C classes, O(|V(G)| C^w)

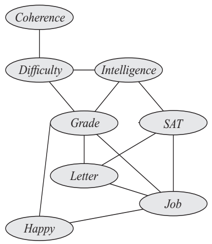
ELIMINATION ACCORDING TO THE ORDER C,D,I,H,G,S,L



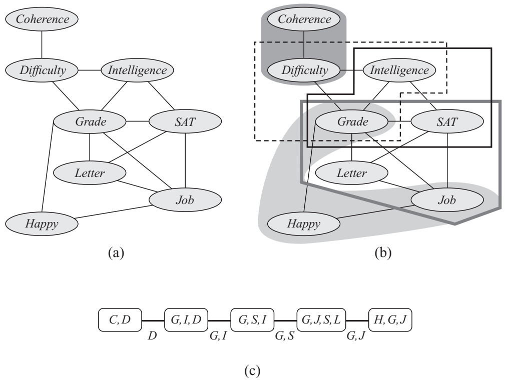
(a) (b) (c)

ELIMINATION ACCORDING TO THE ORDER C,D,I,H,G,S,L



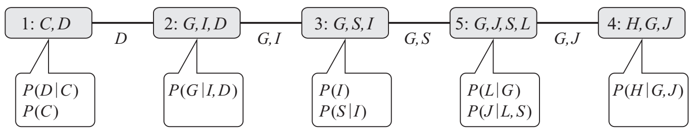


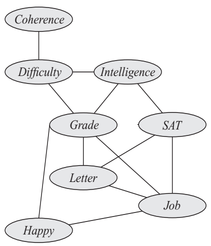
INDUCED GRAPH AND JUNCTION TREE



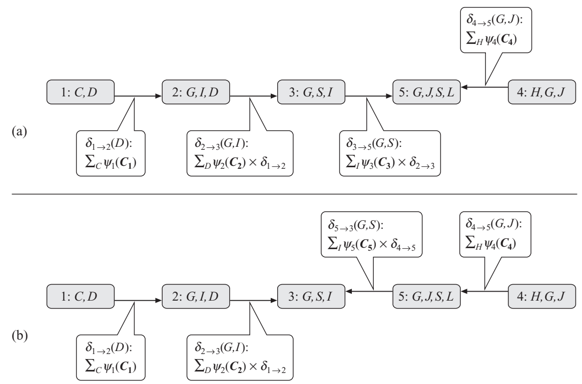
THE END

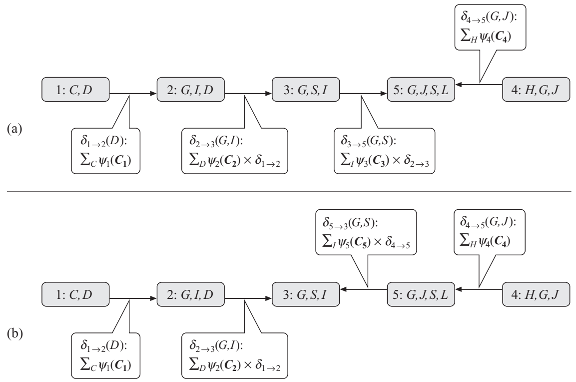
USING THE SIMPLER ELIMINATION & JUNCTION TREE

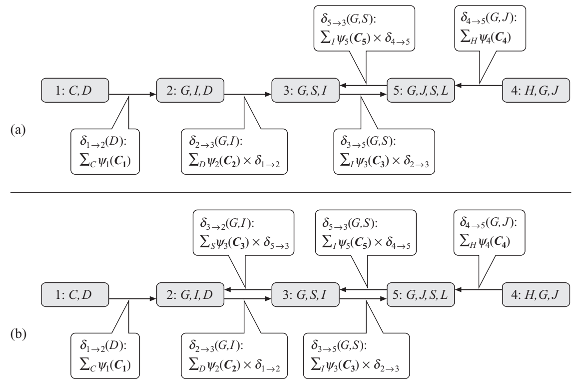




MESSAGE PASSING




(a)
1: C, D
2: G, I, D
3: G, S, I
5: G, J, S, L
4: H, G, J
$\delta_{1\to2}(D)$: $\sum_C \psi_1(\boldsymbol{C_1})$
$\delta_{2\to3}(G,I)$: $\sum_D \psi_2(\boldsymbol{C_2}) \times \delta_{1\to2}$
$\delta_{3\to5}(G,S)$: $\sum_I \psi_3(\boldsymbol{C_3}) \times \delta_{2\to3}$
$\delta_{4\to5}(G,J)$: $\sum_H \psi_4(\boldsymbol{C_4})$
(b)
1: C, D
2: G, I, D
3: G, S, I
5: G, J, S, L
4: H, G, J
$\delta_{1\to2}(D)$: $\sum_C \psi_1(\boldsymbol{C_1})$
$\delta_{2\to3}(G,I)$: $\sum_D \psi_2(\boldsymbol{C_2}) \times \delta_{1\to2}$
$\delta_{5\to3}(G,S)$: $\sum_I \psi_5(\boldsymbol{C_5}) \times \delta_{4\to5}$
$\delta_{4\to5}(G,J)$: $\sum_H \psi_4(\boldsymbol{C_4})$


(a)
1: C, D
2: G, I, D
3: G, S, I
5: G, J, S, L
4: H, G, J
$\delta_{5\to3}(G,S)$: $\sum_I \psi_5(\boldsymbol{C_5}) \times \delta_{4\to5}$
$\delta_{4\to5}(G,J)$: $\sum_H \psi_4(\boldsymbol{C_4})$
$\delta_{1\to2}(D)$: $\sum_C \psi_1(\boldsymbol{C_1})$
$\delta_{2\to3}(G,I)$: $\sum_D \psi_2(\boldsymbol{C_2}) \times \delta_{1\to2}$
$\delta_{3\to5}(G,S)$: $\sum_I \psi_3(\boldsymbol{C_3}) \times \delta_{2\to3}$
(b)
1: C, D
2: G, I, D
3: G, S, I
5: G, J, S, L
4: H, G, J
$\delta_{3\to2}(G,I)$: $\sum_S \psi_3(\boldsymbol{C_3}) \times \delta_{5\to3}$
$\delta_{5\to3}(G,S)$: $\sum_I \psi_5(\boldsymbol{C_5}) \times \delta_{4\to5}$
$\delta_{4\to5}(G,J)$: $\sum_H \psi_4(\boldsymbol{C_4})$
$\delta_{1\to2}(D)$: $\sum_C \psi_1(\boldsymbol{C_1})$
$\delta_{2\to3}(G,I)$: $\sum_D \psi_2(\boldsymbol{C_2}) \times \delta_{1\to2}$
$\delta_{3\to5}(G,S)$: $\sum_I \psi_3(\boldsymbol{C_3}) \times \delta_{2\to3}$