CS6462

Probabilistic and Explainable AI

Lesson 18

Bayesian Networks

*

Construction of Bayesian Networks

by Emil Vassev

March 10, 2025

Building Bayesian Networks

Stages:

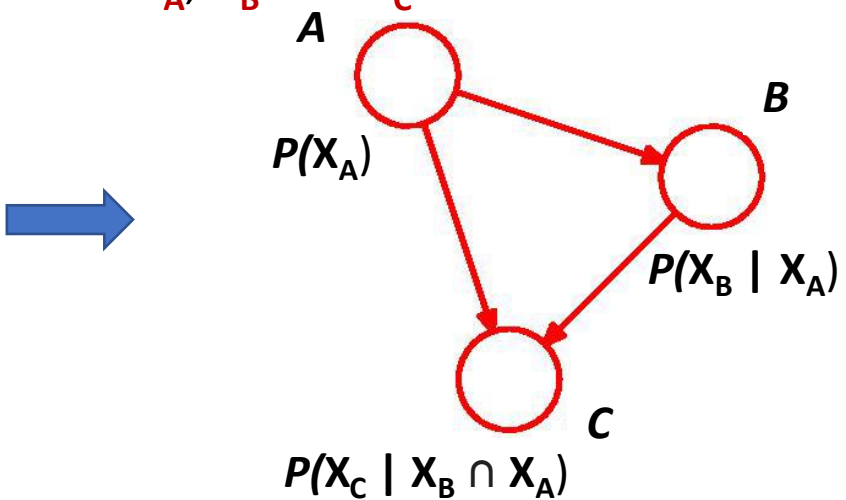
- *Stage 1*: build a directed acyclic graph
  - **V = {A, B, C}, E = {(a,b), (a,c), (b,c)}**
  - every node **V** is associated with a random variable **X**

    V -> X =$\{X_A, X_B, X_C\}$, A -> X_A, B -> X_B, C -> X_C

DAG: $\mathcal{G} = < V, E >$ - nodes V and edges E

- *Stage 2:* assess the conditional probability distribution for every random variable
  - **$P(X_A)$, $P(X_B \mid X_A)$, $P(X_C \mid X_B \cap X_A)$**
- *Stage 3:* assess arbitrary joint distribution over the random variables
  - **$P(X_A \cap X_B \cap X_C) = P(X_A) * P(X_B \mid X_A) * P(X_C \mid X_B \cap X_A)$** holds for any value of **X_A**, **X_B** and **X_C**
- *Stage 4:* represent joint distribution in a simple graphical model:

1) introduce a node for each of the random variables
2) associate each node with the corresponding conditional distribution factor from the joint distribution
3) for each conditional distribution we add directed links to the graph from the nodes corresponding to the variables on which the distribution is conditioned



Example

Bayesian network: Student's Recommendation Letter

- student's grade depends on intelligence and module difficulty
- student asks for a recommendation letter based on the student's grade
- $\boldsymbol{G(V,E)}$: $\boldsymbol{V = \{D}$(module difficulty), $\boldsymbol{I}$(intelligence), $\boldsymbol{G}$(grade), $\boldsymbol{S}$(SAT), $\boldsymbol{L}$(letter)$\boldsymbol{\}}$
- $\mathbf{X = \{X_D, X_I, X_G, X_S, X_L\}}$

  module difficulty: $\mathbf{X_D}$ = {easy, hard}

  intelligence: $\mathbf{X_I}$ = {low, high}

  grade: $\mathbf{X_G}$ = {A, B, C}

  SAT result: $\mathbf{X_S}$ = {high, low}

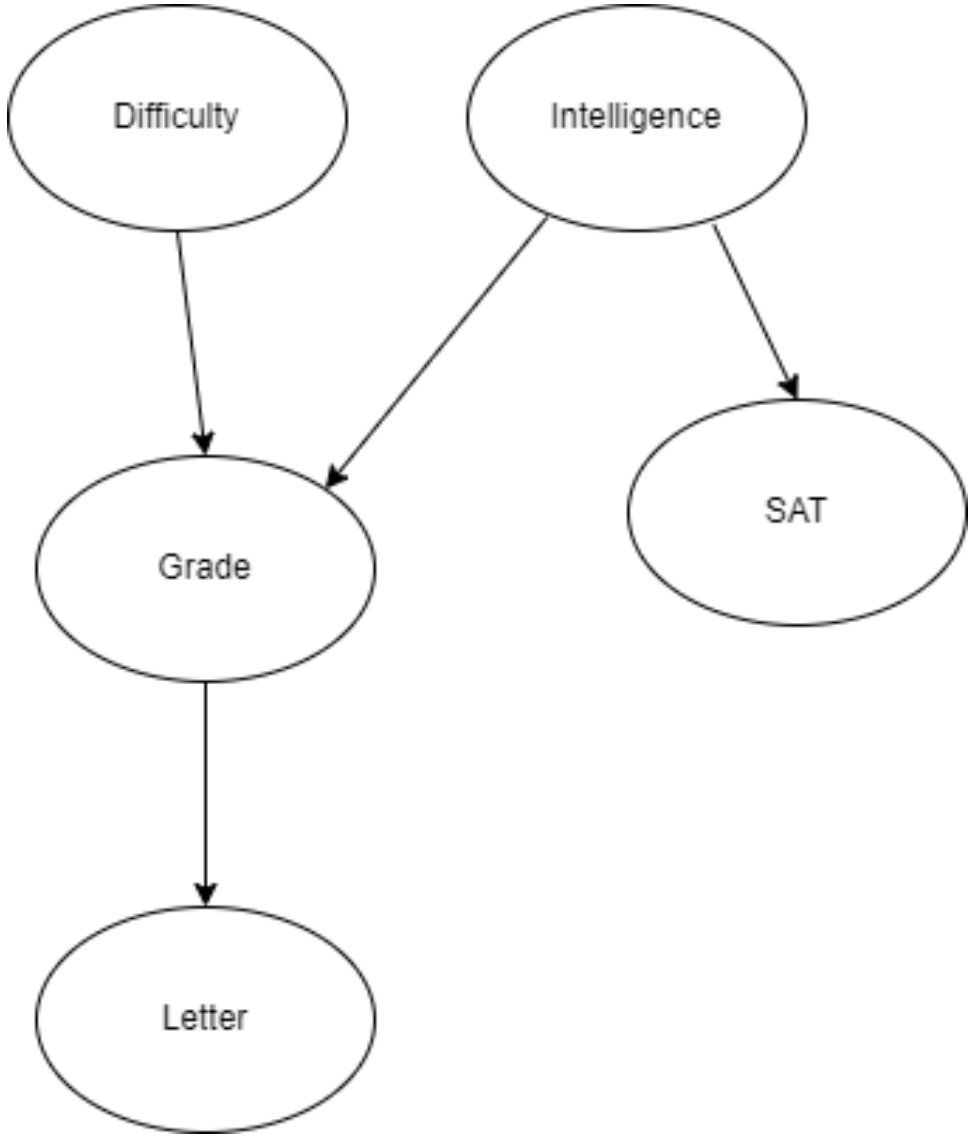
  letter: $\mathbf{X_L}$= {weak, strong}

Example (cont.)

$\mathcal{G}(\boldsymbol{V}, \boldsymbol{E})$: $\boldsymbol{V} = \{\boldsymbol{D}$(module difficulty), $\boldsymbol{I}$(intelligence), $\boldsymbol{G}$(grade), $\boldsymbol{S}$(SAT), $\boldsymbol{L}$(letter)$\}$

- generative sampling process:
- variable's values – selected using distribution that depends on its parents
- **X** – a stochastic function of its parents

X = $\{X_D, X_I, X_G, X_S, X_L\}$
module difficulty: X_D = {easy, hard}
intelligence: X_I = {low, high}
grade: X_G = {A, B, C}
SAT: X_S = {high, low}
letter: X_L = {weak, strong}



Example (cont.)

$\mathbf{X} = \{X_D, X_I, X_G, X_S, X_L\}$
module difficulty: $X_D = \{easy, hard\}$
intelligence: $X_I = \{low, high\}$
grade: $X_G = \{A, B, C\}$
SAT: $X_S = \{low, high\}$
letter: $X_L = \{weak, strong\}$

- local probability models

$P(X_I) = \{0.7, 0.3\}$

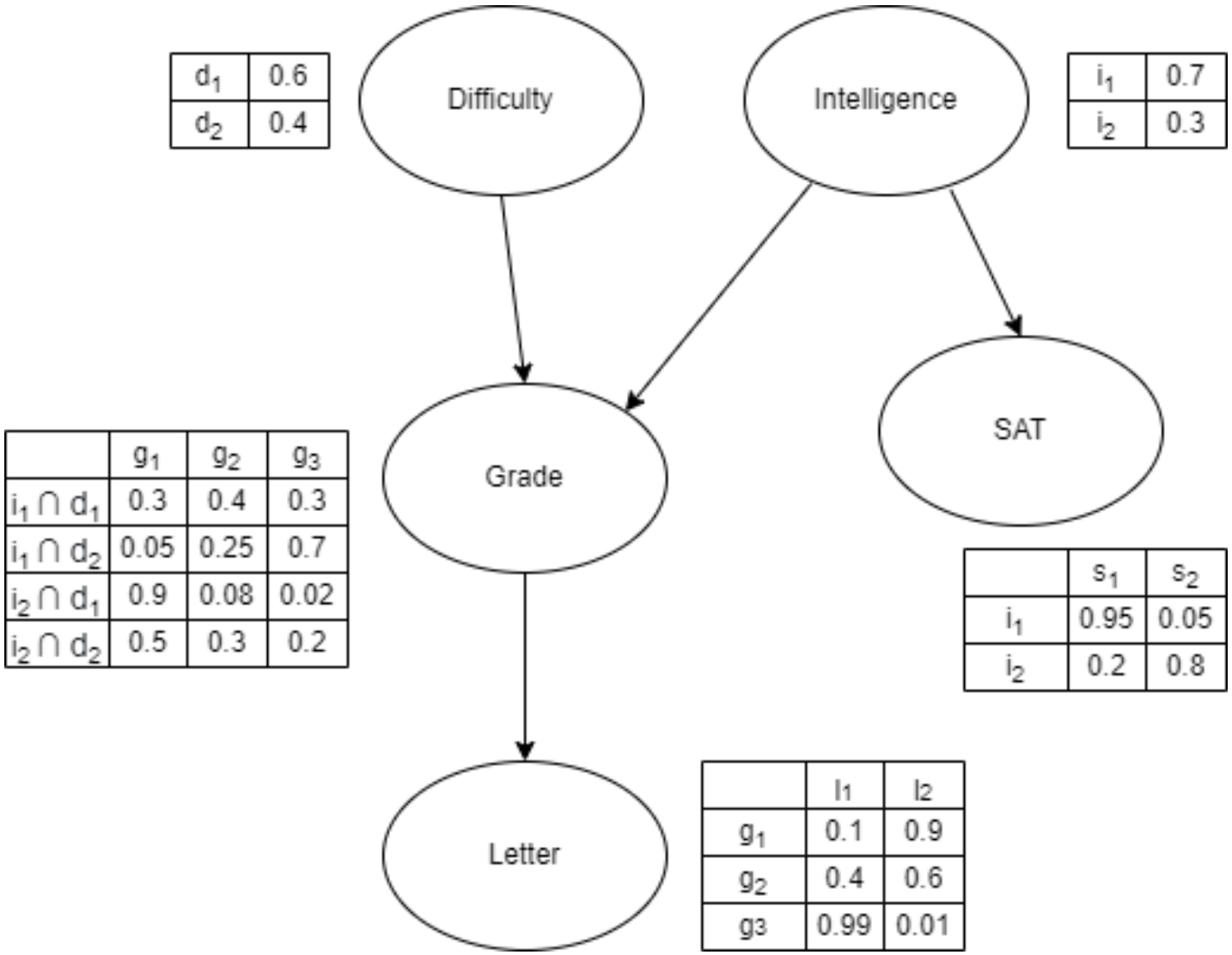
$P(X_D) = \{0.6, 0.4\}$

$P(X_G \mid X_I \cap X_D)$

$P(X_G=A \mid X_I=high \cap X_D=easy) = 0.9$

$P(X_G=B \mid X_I=high \cap X_D=easy) = 0.08$

$P(X_G=C \mid X_I=high \cap X_D=easy) = 0.02$



d_1	0.6
d_2	0.4

i_1	0.7
i_2	0.3

	g_1	g_2	g_3
$i_1 \cap d_1$	0.3	0.4	0.3
$i_1 \cap d_2$	0.05	0.25	0.7
$i_2 \cap d_1$	0.9	0.08	0.02
$i_2 \cap d_2$	0.5	0.3	0.2

	s_1	s_2
i_1	0.95	0.05
i_2	0.2	0.8

	l_1	l_2
g_1	0.1	0.9
g_2	0.4	0.6
g_3	0.99	0.01

Example (cont.)

- local probability models

$P(X_G \mid X_I \cap X_D)$

$P(X_G=A \mid X_I=high \cap X_D=hard) = 0.5$

$P(X_G=B \mid X_I=high \cap X_D=hard) = 0.3$

$P(X_G=C \mid X_I=high \cap X_D=hard) = 0.2$

$P(X_G=A \mid X_I=low \cap X_D=hard) = 0.05$

$P(X_G=B \mid X_I=low \cap X_D=hard) = 0.25$

$P(X_G=C \mid X_I=low \cap X_D=hard) = 0.7$

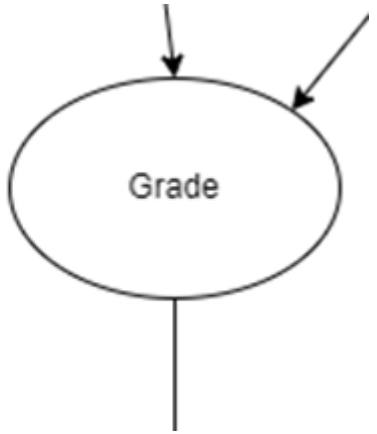
$P(X_G=A \mid X_I=low \cap X_D=easy) = 0.3$

$P(X_G=B \mid X_I=low \cap X_D=easy) = 0.4$

$P(X_G=C \mid X_I=low \cap X_D=easy) = 0.3$

$X = \{X_D, X_I, X_G, X_S, X_L\}$
module difficulty: $X_D = \{easy, hard\}$
intelligence: $X_I = \{low, high\}$
grade: $X_G = \{A, B, C\}$
SAT: $X_S = \{low, high\}$
letter: $X_L = \{weak, strong\}$

	g_1	g_2	g_3
$i_1 \cap d_1$	0.3	0.4	0.3
$i_1 \cap d_2$	0.05	0.25	0.7
$i_2 \cap d_1$	0.9	0.08	0.02
$i_2 \cap d_2$	0.5	0.3	0.2



Example (cont.)

- conditional probability distribution

  $P(X_G \mid X_I \cap X_D)$ – given joint probability of its parents

- marginal distribution – no parents

  $P(X_I) = \{0.7, 0.3\}$

  $P(X_D) = \{0.6, 0.4\}$

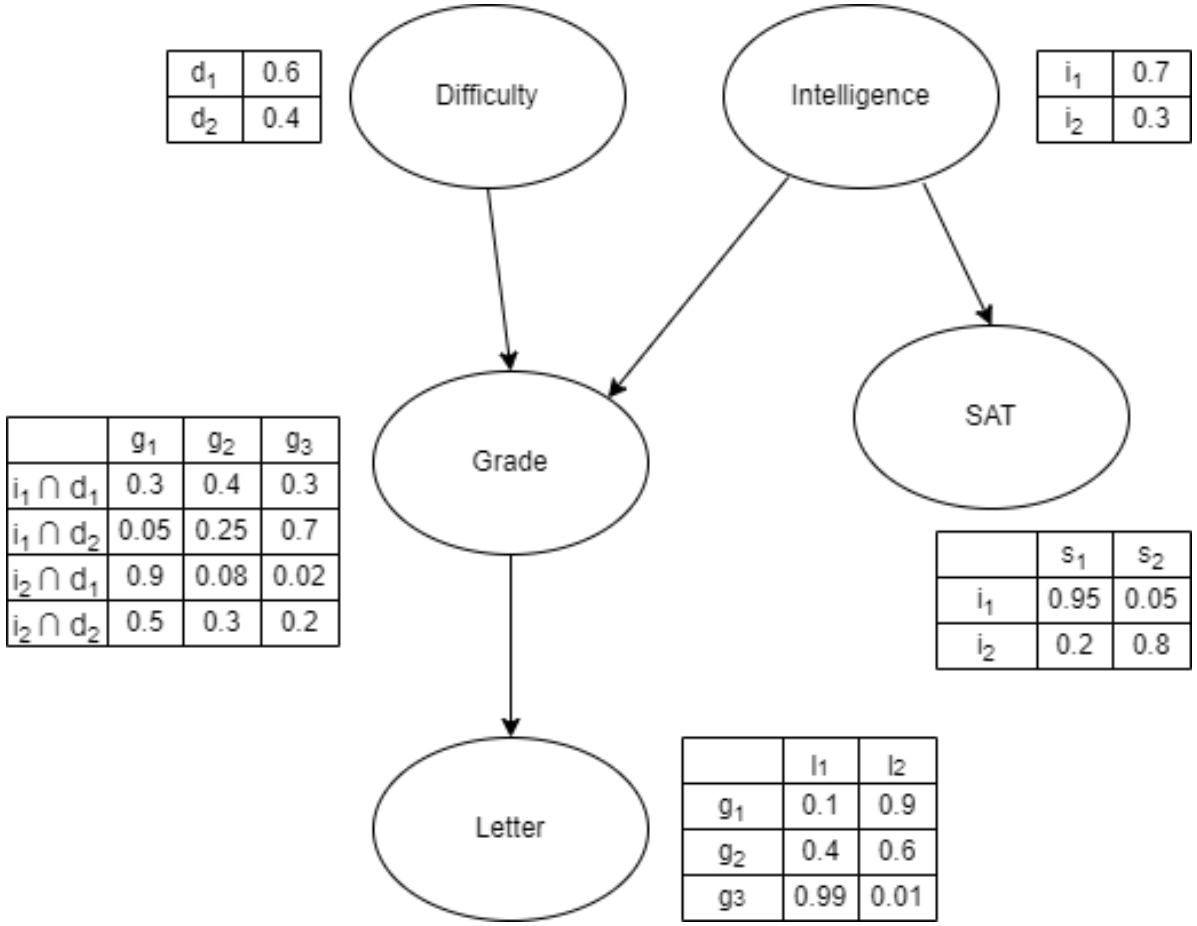
Bayesian Network for

Student's Recommendation Letter:

network structure (DAG) + conditional probability models per node (variable)

joint distribution over random variables?

$X = \{X_D, X_I, X_G, X_S, X_L\}$
module difficulty: $X_D = \{easy, hard\}$
intelligence: $X_I = \{low, high\}$
grade: $X_G = \{A, B, C\}$
SAT: $X_S = \{low, high\}$
letter: $X_L = \{weak, strong\}$



Example (cont.)

$X = \{X_D, X_I, X_G, X_S, X_L\}$
module difficulty: $X_D = \{easy, hard\}$
intelligence: $X_I = \{low, high\}$
grade: $X_G = \{A, B, C\}$
SAT: $X_S = \{low, high\}$
letter: $X_L = \{weak, strong\}$

- Joint distribution

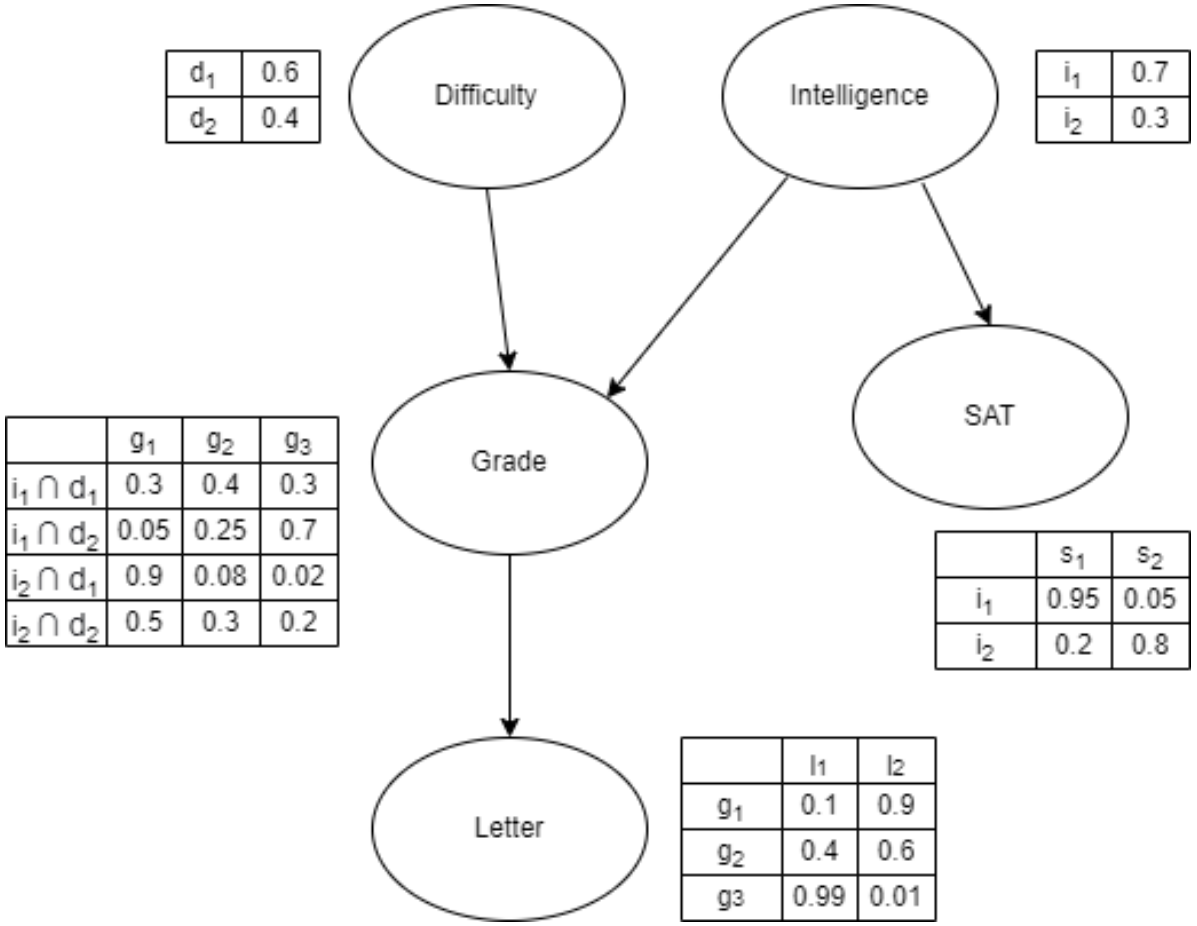
$$P(X_D \cap X_I \cap X_G \cap X_S \cap X_L)$$

$$P(X_D \cap X_I \cap X_G \cap X_S \cap X_L) = P(X_D) * P(X_I | X_D) * P(X_G | X_I \cap X_D) * P(X_S | X_G \cap X_I \cap X_D) * P(X_L | X_S \cap X_G \cap X_I \cap X_D) = P(X_D) * P(X_I) * P(X_G | X_I \cap X_D) * P(X_S | X_I) * P(X_L | X_G)$$

- Example: d_1, i_2, g_1, s_2, l_2 (easy, high, A, high, strong)

$$P(X_D=d_1 \cap X_I=i_2 \cap X_G=g_1 \cap X_S=s_2 \cap X_L=l_2) = P(X_D=d_1) * P(X_I=i_2) * P(X_G=g_1 | X_I=i_2 \cap X_D=d_1) * P(X_S=s_2 | X_I=i_2) * P(X_L=l_2 | X_G=g_1) =$$

$$0.6 * 0.3 * 0.9 * 0.8 * 0.9 = 0.116 = 11.6\%$$



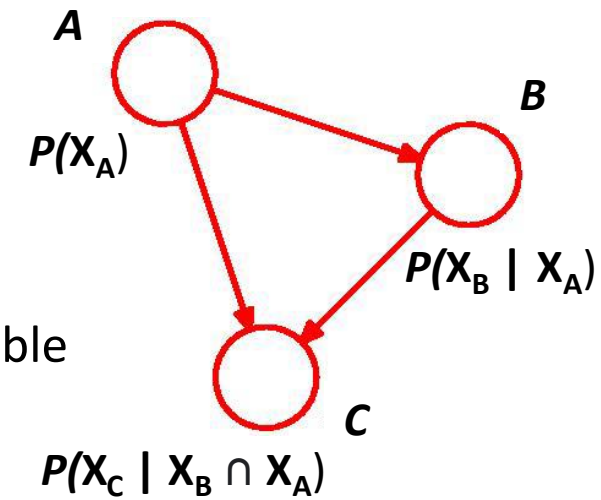
Summary

Bayesian Networks – *Construction of Bayesian Networks*

Stages

- *Stage 1*: build a directed acyclic graph
  - **V = {A, B, C}, E = {(a,b), (a,c), (b,c)}**
  - every node **V** is associated with a random variable **X**

    V -> X =$\{X_A, X_B, X_C\}$, A -> X_A, B -> X_B, C -> X_C
- *Stage 2:* assess the conditional probability distribution for every random variable
  - **$P(X_A)$, $P(X_B \mid X_A)$, $P(X_C \mid X_B \cap X_A)$**
- *Stage 3:* assess arbitrary joint distribution over the random variables
  - **$P(X_A \cap X_B \cap X_C) = P(X_A) * P(X_B \mid X_A) * P(X_C \mid X_B \cap X_A)$** holds for any value of **X_A**, **X_B** and **X_C**
- *Stage 4:* represent joint distribution in a simple graphical model



Next Lesson:

- Bayesian Neural Networks - Bayesian Inference

Thank You!

Questions?