Markov Networks

Probabilistic Graphical Models

Jerónimo Hernández-González

Concepts review

What we have already seen:

- The probabilistic approach to AI
- Factorization, Conditional independence
- What is a probabilistic graphical model
- What is a Bayesian network

Today:

- Factor algebra
- Markov networks

Markov networks

Misconception example

1. Organize four students **in pairs** to work on homework
   1.1 Alice and Bob are friends
   1.2 Bob and Charles study together
   1.3 Charles and Debbie work together but usually disagree
   1.4 Debbie and Alice study together

   1.5 Alice and Charles cannot stand each other
   1.6 Bob and Debbie had a relationship that ended badly
2. *Some* students find out a mistake: the teacher explained something wrong
3. Students transmit this finding to their study partners
4. Probability that all of them are aware of the mistake?

Markov networks

Misconception example

1. Organize four students **in pairs** to work on homework
   1.1 Alice and Bob are friends
   1.2 Bob and Charles study together
   1.3 Charles and Debbie work together but usually disagree
   1.4 Debbie and Alice study together

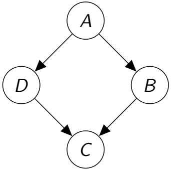
   1.5 Alice and Charles cannot stand each other
   1.6 Bob and Debbie had a relationship that ended badly
2. *Some* students find out a mistake: the teacher explained something wrong
3. Students transmit this finding to their study partners
4. Probability that all of them are aware of the mistake?

$$A \perp\!\!\!\perp C \mid D, B \qquad B \perp\!\!\!\perp D \mid A, C$$

Markov networks

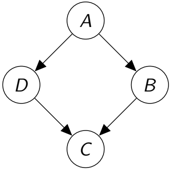
Misconception example

Let us try with Bayesian networks

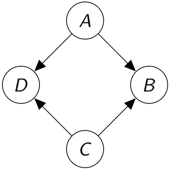


$B \not\perp D \mid A, C$

Let us try with Bayesian networks



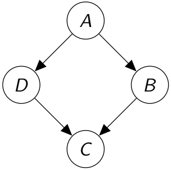
$B \not\perp D \mid A, C$



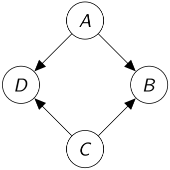
$A \not\perp C \mid D, B$

Markov networks
Misconception example

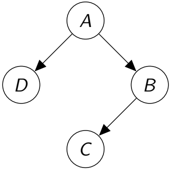
Let us try with Bayesian networks



$B \not\perp\!\!\!\perp D \mid A, C$



$A \not\perp\!\!\!\perp C \mid D, B$

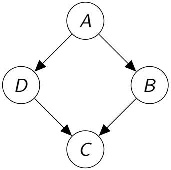


$C \perp\!\!\!\perp D \mid A, B$

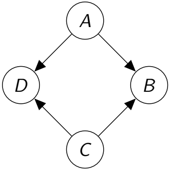
Markov networks

Misconception example

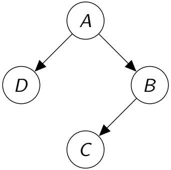
Let us try with Bayesian networks



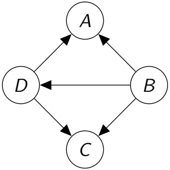
$B \not\perp\!\!\!\perp D \mid A, C$



$A \not\perp\!\!\!\perp C \mid D, B$



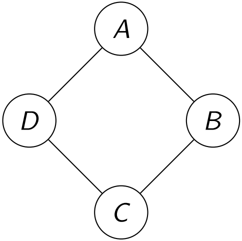
$C \perp\!\!\!\perp D \mid A, B$



$B \not\perp\!\!\!\perp D \mid A, C$

Markov networks

Misconception example

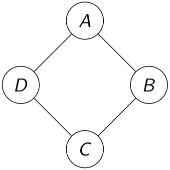


Markov networks vs. Bayesian Networks

Model	Bayesian network	Markov network
Graph	Directed Acyclic	Undirected
Factorization	Distribution per variable (variable and its parents)	Factors over the cliques of the graph
Joint		Requires normalization term (partition function)
Independence	Determined by d-separation in the graph	Determined by separation in the graph
$F \implies I$	Always	Always
$I \implies F$	Always	Only for positive distributions
Markov blanket	Parents, children and parents of children	The neighbors

Markov networks

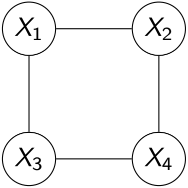
Misconception example

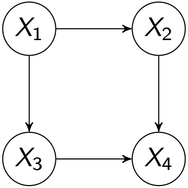


$\phi_1(A,B)$			$\phi_2(B,C)$			$\phi_3(C,D)$			$\phi_4(D,A)$		
a^0	b^0	30	b^0	c^0	100	c^0	d^0	1	d^0	a^0	100
a^0	b^1	5	b^0	c^1	1	c^0	d^1	100	d^0	a^1	1
a^1	b^0	1	b^1	c^0	1	c^1	d^0	100	d^1	a^0	1
a^1	b^1	10	b^1	c^1	100	c^1	d^1	1	d^1	a^1	100

Markov networks vs. Bayesian networks

1. Graph





2. Factorization

$$P(\boldsymbol{x}) = \frac{1}{Z} \prod_{i=1}^{f} \phi_i(\boldsymbol{x}_{\phi_i})$$

$$P(\boldsymbol{x}) = \prod_{i=1}^{v} p(\boldsymbol{x}_i \mid \boldsymbol{pa}_i)$$

There is no partition function

3. Factors

X_4	X_2	$\phi(X_4, X_2)$
a	a	3,3
b	a	13,2
a	b	18,0
b	b	12,0

Factor potentials

X_4	X_2	$p(X_4 \vert X_2)$
a	a	0,25
b	a	0,75
a	b	0,66
b	b	0,33

Conditional prob. distributions

Markov networks vs. Bayesian Networks

Distributions and factors

- Distributions:
  - Joint distribution
  - Common operators:
    - Conditioning
    - Marginalization
- Factors:
  - How are they different from distributions?
  - Common operators:
    - Product
    - Reduction
    - Marginalization
    - Normalization

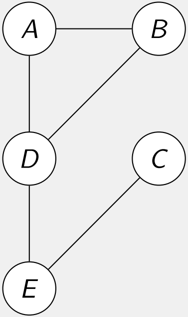
Markov Networks

Probabilistic Graphical Models

Jerónimo Hernández-González

Clique or complete subgraph

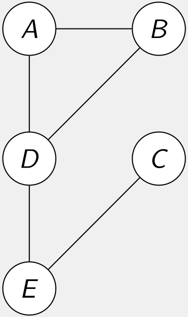
Subgraph where every two vertices are connected to each other



- 1-vertex cliques:
  $C_1 = \{A\}$, $C_2 = \{B\}$, $C_3 = \{C\}$,
  $C_4 = \{D\}$, $C_5 = \{E\}$
- 2-vertex cliques:
  $C_6 = \{A, B\}$, $C_7 = \{A, D\}$,
  $C_8 = \{B, D\}$, $C_9 = \{D, E\}$,
  $C_{10} = \{E, C\}$
- 3-vertex cliques:
  $C_{11} = \{A, B, D\}$

Clique or complete subgraph

Subgraph where every two vertices are connected to each other



- 1-vertex cliques:
  $C_1 = \{A\}$, $C_2 = \{B\}$, $C_3 = \{C\}$,
  $C_4 = \{D\}$, $C_5 = \{E\}$
- 2-vertex cliques:
  $C_6 = \{A, B\}$, $C_7 = \{A, D\}$,
  $C_8 = \{B, D\}$, $C_9 = \{D, E\}$,
  $C_{10} = \{E, C\}$
- 3-vertex cliques:
  $C_{11} = \{A, B, D\}$

Maximal clique

There is no any other larger clique that contains it

Markov networks

Factorization

Let $\mathcal{H}$ be an undirected graph. A distribution $P(\boldsymbol{X})$ factorizes according to $\mathcal{H}$ iff

$$P(\boldsymbol{x}) = \frac{1}{Z} \prod_{i=1}^{f} \phi_i(\boldsymbol{x}_{\phi_i})$$

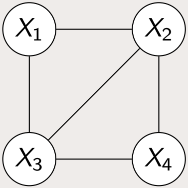
where, for each ϕ_i, there is an edge in $\mathcal{H}$ between each pair of variables in $Scope(\phi_i)$ (i.e., $Scope(\phi_i)$ forms a clique in $\mathcal{H}$)

and $Z = \sum_x \prod_{i=1}^{n} \phi_i(x_{\phi_i})$ is the normalization constant, known as *partition function*.

Exercise

Factorization in MNs

Which of these factorizations is valid for the following graph?



$$P(X) \propto \phi_1(X_4)\phi_2(X_1, X_2, X_3)$$

$$P(X) \propto \phi_1(X_1, X_2, X_4)$$

$$P(X) \propto \phi_1(X_1, X_2, X_3)\phi_2(X_2, X_3, X_4)$$

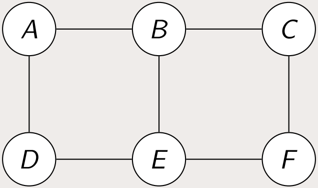
$$P(X) \propto \phi_1(X_1)\phi_2(X_1, X_2)\phi_3(X_3)\phi_4(X_3, X_4)$$

$$P(X) \propto \phi_1(X_1, X_2)\phi_2(X_1, X_3)\phi_3(X_3, X_4)\phi_4(X_2, X_4)\phi_5(X_2, X_3)$$

Exercise

Factorization in MNs

Which of the following sets of factors could factorize over this undirected graph?



a) $\phi(A), \phi(B), \phi(C), \phi(D), \phi(E), \phi(F)$

b) $\phi(A, B, D), \phi(A, B), \phi(C, D, E), \phi(E, F), \phi(F)$

c) $\phi(A, B, D, C), \phi(C, D, E, F)$

d) $\phi(A, B), \phi(C, D), \phi(E, F)$

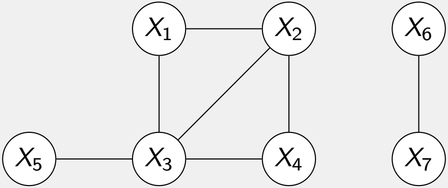
Markov Networks

Probabilistic Graphical Models

Jerónimo Hernández-González

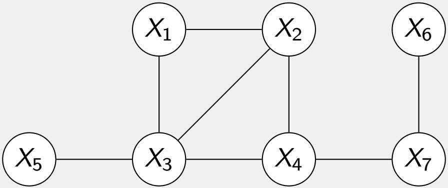
Independence

Given a Markov network structure $\mathcal{H}$, two variables X and Y are independent in $\mathcal{H}$ if there is no path between them in $\mathcal{H}$



Conditional independence

Given a Markov network structure $\mathcal{H}$, two variables X and Y are independent given a third variable Z if Z separates X from Y in $\mathcal{H}$



Intuitively, probabilistic influence flow is stopped at observed nodes

Flow of probabilistic influence

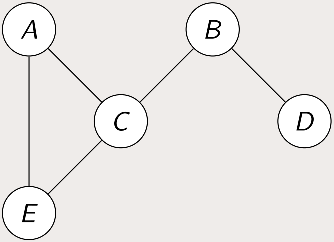
Active trail

- Let $\mathcal{H}$ be an undirected graph
- Let $X_1 \rightleftharpoons \ldots \rightleftharpoons X_m$ be a trail in $\mathcal{H}$
- A trail is active given a set of observed variables $\boldsymbol{Z}$ if no node along the trail is in $\boldsymbol{Z}$

Exercise

Independence in MNs

Which pairs of variables are independent in this network?



a) $A \perp\!\!\!\perp C$

b) $C \perp\!\!\!\perp D$

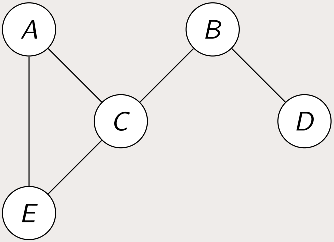
c) $D \perp\!\!\!\perp E$

d) $A \perp\!\!\!\perp D$

Exercise

Independence in MNs

Which conditional independence statements hold in this network?



a) $A \perp\!\!\!\perp C | E$

b) $C \perp\!\!\!\perp D | B$

c) $D \perp\!\!\!\perp E | C$

d) $A \perp\!\!\!\perp D | E$

Markov network vs. Bayesian network

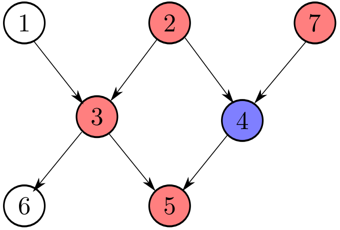
Markov blanket

Definition

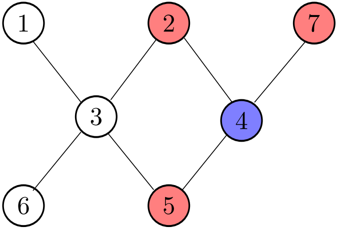
Given X_j and $\boldsymbol{Mb}_j$, for any set of variables $\boldsymbol{X}_c \subseteq (\boldsymbol{V} \setminus \boldsymbol{Mb}_j)$:

$$\boldsymbol{X}_c \perp\!\!\!\perp X_j \mid \boldsymbol{Mb}_j$$

Parents, children and parents of the children

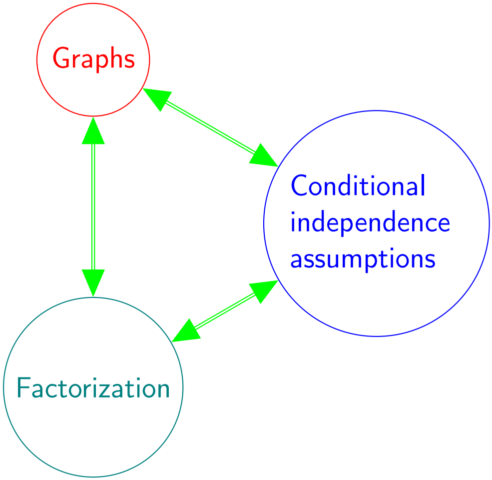


Neighbors



Markov networks

Connection between the three views



- Given a Markov network structure $\mathcal{H}$, and a probability distribution P that factorizes over $\mathcal{H}$, then P satisfies the independencies that hold in $\mathcal{H}$.

$$P \text{ factorizes} \implies P \text{ satisfies independencies}$$

- Given a Markov network structure $\mathcal{H}$, and a *positive* probability distribution P that satisfies the independencies that hold in $\mathcal{H}$, then P factorizes over $\mathcal{H}$.

$$\left.\begin{array}{ll} P & \text{satisfies independencies} \\ P & \text{positive} \end{array}\right\} \implies P \text{ factorizes}$$

Exercise

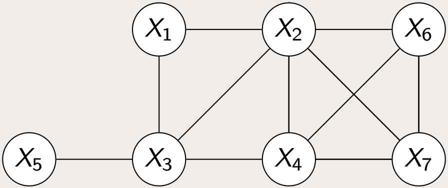
Let P be a distribution such that

$$P(X_1, X_2, X_3, X_4) \propto \phi_1(X_1)\phi_2(X_3, X_4)\phi_3(X_1, X_4)\phi_4(X_4, X_2)$$

- $X_1 \perp\!\!\!\perp X_3 \mid X_4$?
- $X_2 \perp\!\!\!\perp X_3 \mid X_1$?
- $X_2, X_3 \perp\!\!\!\perp X_1$?

Exercise

Given that P is positive and satisfies the independencies in:



Provide a factorization for P.

Markov Networks

Probabilistic Graphical Models

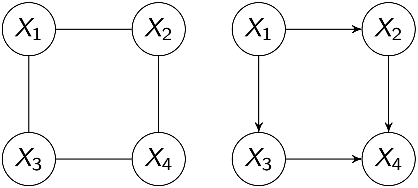
Jerónimo Hernández-González

Markov networks summary

Model name	Markov network
Graph type	Undirected
Factorization	Over the cliques of the graph
Probability distr.	Requires a normalization term (partition funct.)
Independence	Determined by separation in the graph
$F \implies I$	Always
$I \implies F$	Only for positive distributions
Markov blanket	The neighbors

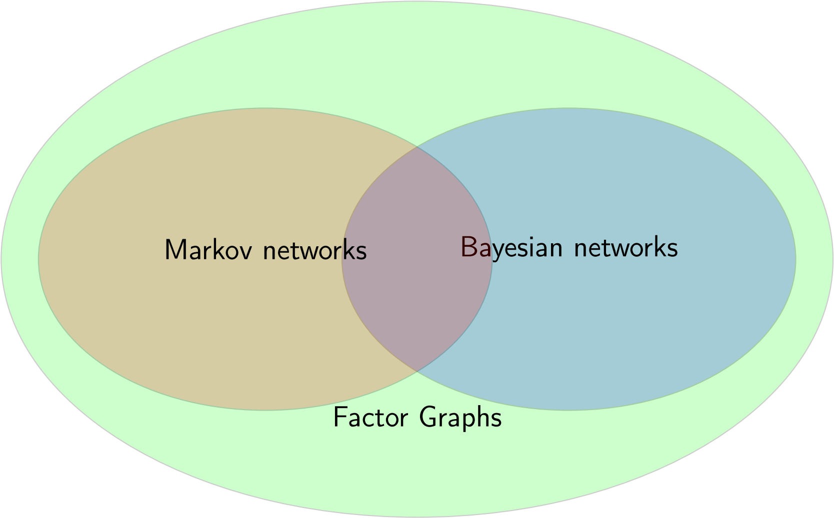
Expressiveness

Models represented by Markov networks and Bayesian networks



Expressiveness

Models represented by Markov networks and Bayesian networks



** The intersection covers the decomposable models: BNs without V-structures

Exercise

Storing a Markov network

Let $P(X)$ be a probability distribution that factorizes according to a Markov network $\mathcal{H}$. How much memory do we need?

Markov Networks

Probabilistic Graphical Models

Jerónimo Hernández-González

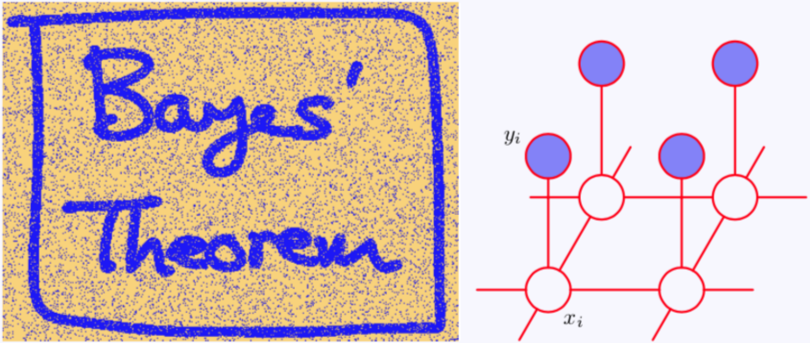
Markov networks

Uses

from Christopher Burger, Christian Schuler and Stefan Harmeling. CVPR 2012.

Markov networks

Uses



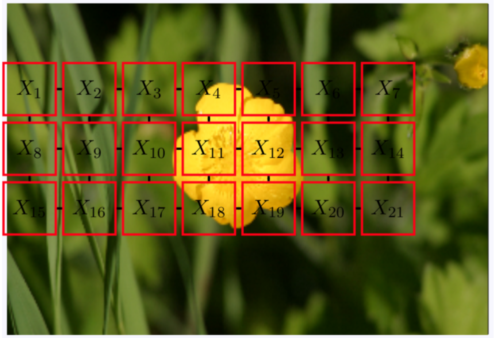
Noisy image Markov Network

Markov networks

Uses

Markov networks

Uses



Markov Networks

Probabilistic Graphical Models

Jerónimo Hernández-González

Factor algebra

- $\boldsymbol{X} = (X_1, \ldots, X_N)$ is a set of variables
- A factor ϕ is a **function**,

$$\phi : Val(\boldsymbol{X}_\phi) \rightarrow \mathbb{R},$$

over a subset of variables. This set of variables is called *scope*:

$$Scope(\phi) = \boldsymbol{X}_\phi \subseteq \boldsymbol{X}$$

- Each variable X_i has its own set of possible values, $Val(X_i)$ or Ω_{X_i}.
- Similarly, for a set of variables $\boldsymbol{X}_\phi$, we define $Val(\boldsymbol{X}_\phi)$ or $\Omega_{\boldsymbol{X}_\phi}$ as the **cartesian product** of the sets of possible values of the variables in $\boldsymbol{X}_\phi$.

** Remember: assuming that random variables are **discrete**

Exercise

Factor Scope

Let $\phi(c, e)$ be a factor in a graphical model, where c is a value of C and e is a value of E. Which is the scope of ϕ?

a) C, E

b) A, C, E

c) A, B, C, E

d) C

Factor Algebra

Examples

X	Y	ϕ
0	0	3
0	1	2
1	0	4
1	1	1

Y	Z	ψ
0	0	5
0	1	4
0	2	1
1	0	2
1	1	0
1	2	6

Operations

- Product
- Reduction
- Marginalization
- Normalization

Factor algebra

Product

Product of factors

Given two factors ϕ and ψ, their product $\phi \times \psi$ is a new factor whose scope is the union of the scopes of ϕ and ψ ($\Omega_{X_\phi} \cup \Omega_{X_\psi}$) and whose value is the product of ϕ and ψ.

Product of factors

Given two factors ϕ and ψ, their product $\phi \times \psi$ is a new factor whose scope is the union of the scopes of ϕ and ψ ($\Omega_{X_\phi} \cup \Omega_{X_\psi}$) and whose value is the product of ϕ and ψ.

X	Y	ϕ
0	0	3
0	1	2
1	0	4
1	1	1

Y	Z	ψ
0	0	5
0	1	4
0	2	1
1	0	2
1	1	0
1	2	6

Scope ($\Omega_{X_\phi} \cup \Omega_{X_\psi}$) of $\phi \times \psi$?

Product of factors

Given two factors ϕ and ψ, their product $\phi \times \psi$ is a new factor whose scope is the union of the scopes of ϕ and ψ ($\Omega_{X_\phi} \cup \Omega_{X_\psi}$) and whose value is the product of ϕ and ψ.

X	Y	ϕ
0	0	3
0	1	2
1	0	4
1	1	1

Y	Z	ψ
0	0	5
0	1	4
0	2	1
1	0	2
1	1	0
1	2	6

X	Y	Z	$\phi \times \psi$
0	0	0	
0	0	1	
0	0	2	
0	1	0	
0	1	1	
0	1	2	
1	0	0	
1	0	1	
1	0	2	
1	1	0	
1	1	1	
1	1	2	

Product of factors

Given two factors ϕ and ψ, their product $\phi \times \psi$ is a new factor whose scope is the union of the scopes of ϕ and ψ ($\Omega_{X_\phi} \cup \Omega_{X_\psi}$) and whose value is the product of ϕ and ψ.

X	Y	ϕ
0	0	3
0	1	2
1	0	4
1	1	1

Y	Z	ψ
0	0	5
0	1	4
0	2	1
1	0	2
1	1	0
1	2	6

X	Y	Z	$\phi \times \psi$
0	0	0	15
0	0	1	12
0	0	2	3
0	1	0	4
0	1	1	0
0	1	2	12
1	0	0	20
1	0	1	16
1	0	2	4
1	1	0	2
1	1	1	0
1	1	2	6

Properties of the product

- The product of factors is **commutative**

$$\phi \times \psi = \psi \times \phi$$

- The product of factors is **associative**

$$(\phi \times \psi) \times \xi = \psi \times (\phi \times \xi)$$

- For each scope $\boldsymbol{U} \subseteq \boldsymbol{X}$, there is a **neutral element** that assigns $\phi(\mathbf{u}) = 1$

Reduction of a factor

The reduction of a factor ϕ for an assignment of values $\boldsymbol{U} = \boldsymbol{u}$ is a new factor $\phi[\boldsymbol{u}]$ whose scope is $\boldsymbol{V} = \boldsymbol{X}_\phi \backslash \boldsymbol{U}$ and whose value for the assignment $\boldsymbol{V} = \boldsymbol{v}$, $\phi[\boldsymbol{u}](\boldsymbol{v})$, is the value of ϕ for the joint assignment of $\boldsymbol{u}$ and $\boldsymbol{v}$, $\phi[\boldsymbol{u}](\boldsymbol{v}) = \phi(\boldsymbol{u}, \boldsymbol{v})$.

Reduction of a factor

The reduction of a factor ϕ for an assignment of values $\boldsymbol{U} = \boldsymbol{u}$ is a new factor $\phi[\boldsymbol{u}]$ whose scope is $\boldsymbol{V} = \boldsymbol{X}_\phi \backslash \boldsymbol{U}$ and whose value for the assignment $\boldsymbol{V} = \boldsymbol{v}$, $\phi[\boldsymbol{u}](\boldsymbol{v})$, is the value of ϕ for the joint assignment of $\boldsymbol{u}$ and $\boldsymbol{v}$, $\phi[\boldsymbol{u}](\boldsymbol{v}) = \phi(\boldsymbol{u}, \boldsymbol{v})$.

X	Y	Z	ϕ
0	0	0	4
0	0	1	3
0	0	2	5
0	1	0	11
0	1	1	2
0	1	2	1
1	0	0	4
1	0	1	5
1	0	2	12
1	1	0	4
1	1	1	1
1	1	2	9

Scope of $\phi[X = 0]$?

Scope of $\phi[Y = 1, Z = 2]$?

Reduction of a factor

The reduction of a factor ϕ for an assignment of values $\boldsymbol{U} = \boldsymbol{u}$ is a new factor $\phi[\boldsymbol{u}]$ whose scope is $\boldsymbol{V} = \boldsymbol{X}_\phi \backslash \boldsymbol{U}$ and whose value for the assignment $\boldsymbol{V} = \boldsymbol{v}$, $\phi[\boldsymbol{u}](\boldsymbol{v})$, is the value of ϕ for the joint assignment of $\boldsymbol{u}$ and $\boldsymbol{v}$, $\phi[\boldsymbol{u}](\boldsymbol{v}) = \phi(\boldsymbol{u}, \boldsymbol{v})$.

X	Y	Z	ϕ
0	0	0	4
0	0	1	3
0	0	2	5
0	1	0	11
0	1	1	2
0	1	2	1
1	0	0	4
1	0	1	5
1	0	2	12
1	1	0	4
1	1	1	1
1	1	2	9

Y	Z	$\phi[X = 0]$
0	0	
0	1	
0	2	
1	0	
1	1	
1	2	

X	$\phi[Y = 1, Z = 2]$
0	
1	

Reduction of a factor

The reduction of a factor ϕ for an assignment of values $\boldsymbol{U} = \boldsymbol{u}$ is a new factor $\phi[\boldsymbol{u}]$ whose scope is $\boldsymbol{V} = \boldsymbol{X}_{\phi} \setminus \boldsymbol{U}$ and whose value for the assignment $\boldsymbol{V} = \boldsymbol{v}$, $\phi[\boldsymbol{u}](\boldsymbol{v})$, is the value of ϕ for the joint assignment of $\boldsymbol{u}$ and $\boldsymbol{v}$, $\phi[\boldsymbol{u}](\boldsymbol{v}) = \phi(\boldsymbol{u}, \boldsymbol{v})$.

X	Y	Z	ϕ
0	0	0	4
0	0	1	3
0	0	2	5
0	1	0	11
0	1	1	2
0	1	2	1
1	0	0	4
1	0	1	5
1	0	2	12
1	1	0	4
1	1	1	1
1	1	2	9

Y	Z	$\phi[X = 0]$
0	0	4
0	1	3
0	2	5
1	0	11
1	1	2
1	2	1

X	$\phi[Y = 1, Z = 2]$
0	1
1	9

Relationship between reduction and product of factors

Let ϕ_1 and ϕ_2 be two factors, and $\boldsymbol{U} = \boldsymbol{u}$ an assignment of values to variables:

$$(\phi_1 \times \phi_2)[\boldsymbol{U} = \boldsymbol{u}] = \phi_1[\boldsymbol{U} = \boldsymbol{u}] \times \phi_2[\boldsymbol{U} = \boldsymbol{u}]$$

Relationship between reduction and product of factors

Let ϕ_1 and ϕ_2 be two factors, and $\boldsymbol{U} = \boldsymbol{u}$ an assignment of values to variables:

$$(\phi_1 \times \phi_2)[\boldsymbol{U} = \boldsymbol{u}] = \phi_1[\boldsymbol{U} = \boldsymbol{u}] \times \phi_2[\boldsymbol{U} = \boldsymbol{u}]$$

This rule can help us reducing work!

$$p(E = e) \quad \text{or} \quad p(X \mid E = e)$$

Factor algebra

Marginalization

Marginal

Given a factor ϕ and a set of variables $\boldsymbol{V}$ to remove, the **marginal** $\sum_{\boldsymbol{V}} \phi$ is a factor ψ with scope $\boldsymbol{U} = \boldsymbol{X}_\phi \setminus \boldsymbol{V}$, defined by
$\psi(\boldsymbol{u}) = \sum_{\boldsymbol{v}} \phi(\boldsymbol{u}, \boldsymbol{v})$ **sometimes written as $\phi^{\downarrow \boldsymbol{U}}$

Factor algebra

Marginalization

Marginal

Given a factor ϕ and a set of variables $\boldsymbol{V}$ to remove, the **marginal** $\sum_{\boldsymbol{V}} \phi$ is a factor ψ with scope $\boldsymbol{U} = \boldsymbol{X}_\phi \setminus \boldsymbol{V}$, defined by
$\psi(\boldsymbol{u}) = \sum_{\boldsymbol{v}} \phi(\boldsymbol{u}, \boldsymbol{v})$ **sometimes written as $\phi^{\downarrow \boldsymbol{U}}$

X	Y	Z	ϕ
0	0	0	4
0	0	1	3
0	0	2	5
0	1	0	11
0	1	1	2
0	1	2	1
1	0	0	4
1	0	1	5
1	0	2	12
1	1	0	4
1	1	1	1
1	1	2	9

Scope of $\sum_Y \phi$?

Scope of $\sum_{Y,Z} \phi$?

Factor algebra

Marginalization

Marginal

Given a factor ϕ and a set of variables $\boldsymbol{V}$ to remove, the **marginal** $\sum_{\boldsymbol{V}} \phi$ is a factor ψ with scope $\boldsymbol{U} = \boldsymbol{X}_\phi \setminus \boldsymbol{V}$, defined by $\psi(\boldsymbol{u}) = \sum_{\boldsymbol{v}} \phi(\boldsymbol{u}, \boldsymbol{v})$ **sometimes written as $\phi^{\downarrow \boldsymbol{U}}$

X	Y	Z	ϕ
0	0	0	4
0	0	1	3
0	0	2	5
0	1	0	11
0	1	1	2
0	1	2	1
1	0	0	4
1	0	1	5
1	0	2	12
1	1	0	4
1	1	1	1
1	1	2	9

X	Z	$\sum_Y \phi$
0	0	
0	1	
0	2	
1	0	
1	1	
1	2	

X	$\sum_{Y,Z} \phi$
0	
1	

Factor algebra

Marginalization

Marginal

Given a factor ϕ and a set of variables $\boldsymbol{V}$ to remove, the **marginal** $\sum_{\boldsymbol{V}} \phi$ is a factor ψ with scope $\boldsymbol{U} = \boldsymbol{X}_\phi \setminus \boldsymbol{V}$, defined by $\psi(\boldsymbol{u}) = \sum_{\boldsymbol{v}} \phi(\boldsymbol{u}, \boldsymbol{v})$

**sometimes written as $\phi^{\downarrow \boldsymbol{U}}$

X	Y	Z	ϕ
0	0	0	4
0	0	1	3
0	0	2	5
0	1	0	11
0	1	1	2
0	1	2	1
1	0	0	4
1	0	1	5
1	0	2	12
1	1	0	4
1	1	1	1
1	1	2	9

X	Z	$\sum_Y \phi$
0	0	15
0	1	5
0	2	6
1	0	8
1	1	6
1	2	21

X	$\sum_{Y,Z} \phi$
0	26
1	35

Factor algebra

Product and marginalization

Relationship between marginalization and product of factors

Let ϕ_1 and ϕ_2 be two factors, if $X \notin Scope(\phi_1)$:

$$\sum_X (\phi_1 \times \phi_2) = \phi_1 \times \sum_X \phi_2$$

Factor algebra

Product and marginalization

Relationship between marginalization and product of factors

Let ϕ_1 and ϕ_2 be two factors, if $X \notin Scope(\phi_1)$:

$$\sum_X (\phi_1 \times \phi_2) = \phi_1 \times \sum_X \phi_2$$

This rule can help us reducing work!

$$p(X) = \frac{1}{\theta} \sum_{V \setminus X} \phi_1 \times \cdots \times \phi_f$$

This rule can help us reducing work!

The order of summation matters!

This rule can help us reducing work!

The order of summation matters!

Whenever we have to assess a factor algebra expression:

1. For each reduction operation:
   1.1 Move it towards the affected factors and do them
2. For each marginalization operation:
   2.1 Select a good ordering of the variables
   2.2 Move it towards the affected factors
   2.3 Perform marginalization operations "from the inside out"

Exercise

Reduction and marginalization

X	Y	Z	ϕ
0	0	0	4
0	0	1	3
0	0	2	5
0	1	0	11
0	1	1	2
0	1	2	1
1	0	0	4
1	0	1	5
1	0	2	12
1	1	0	4
1	1	1	1
1	1	2	9

$$p(X|Y=1)$$

$$p(Z|X=0)$$

Exercise

Reduction and marginalization

Let $\{A, B, C, D\}$ be a set of binary variables, where

$$p(A, B, C, D) = \frac{1}{\theta}\phi_1(A, B)\phi_2(B, C)\phi_3(C, D)\phi_4(D, A)$$

- $p(A, C|B = b)$
- $p(A|B = b)$
- $p(D|C = c)$
- $p(B|D = d)$

Exercise

Reduction and marginalization

$\phi_1(A,B)$			$\phi_2(B,C)$			$\phi_3(C,D)$			$\phi_4(D,A)$		
a^0	b^0	30	b^0	c^0	100	c^0	d^0	1	d^0	a^0	100
a^0	b^1	5	b^0	c^1	1	c^0	d^1	100	d^0	a^1	1
a^1	b^0	1	b^1	c^0	1	c^1	d^0	100	d^1	a^0	1
a^1	b^1	10	b^1	c^1	100	c^1	d^1	1	d^1	a^1	100

$$p(A,B,C,D) = \frac{1}{\theta}\phi_1(A,B)\phi_2(B,C)\phi_3(C,D)\phi_4(D,A)$$

$p(D|C=c^1) =$

Marginal

Given a factor ϕ, its normalization

$$Norm(\phi)(\boldsymbol{x}) = \frac{\phi(\boldsymbol{x})}{Z_\phi}$$

where $Z_\phi = \sum_{\boldsymbol{x}} \phi(\boldsymbol{x})$

Factor algebra

Normalization

Marginal

Given a factor ϕ, its normalization

$$Norm(\phi)(\boldsymbol{x}) = \frac{\phi(\boldsymbol{x})}{Z_\phi}$$

where $Z_\phi = \sum_{\boldsymbol{x}} \phi(\boldsymbol{x})$

Y	Z	ψ	$Norm(\psi)$
0	0	2	
0	1	5	
0	2	2	
1	0	4	
1	1	3	
1	2	1	

Z	ϕ	$Norm(\phi)$
0	8	
1	2	

Marginal

Given a factor ϕ, its normalization

$$Norm(\phi)(\boldsymbol{x}) = \frac{\phi(\boldsymbol{x})}{Z_\phi}$$

where $Z_\phi = \sum_{\boldsymbol{x}} \phi(\boldsymbol{x})$

Y	Z	ψ	$Norm(\psi)$
0	0	2	2/17
0	1	5	5/17
0	2	2	2/17
1	0	4	4/17
1	1	3	3/17
1	2	1	1/17

Z	ϕ	$Norm(\phi)$
0	8	8/10=0.8
1	2	2/10=0.2

Factor algebra
Summary

- Product: $\phi_1 \times \phi_2$
- Reduction: $\phi[\boldsymbol{E} = \boldsymbol{e}]$
- Marginalization: $\sum_X \phi$
- Normalization: $Norm(\phi)$
- Product and reduction interaction:

$$(\phi_1 \times \phi_2)\,[\boldsymbol{U} = \boldsymbol{u}] = \phi_1[\boldsymbol{U} = \boldsymbol{u}] \times \phi_2[\boldsymbol{U} = \boldsymbol{u}]$$

- Product and marginalization interaction:

$$\sum_X (\phi_1 \times \phi_2) = \phi_1 \sum_X \phi_2$$

where $X \notin Scope(\phi_1)$

Exercise

Factors in Markov Network

Let $\phi_1(A, B)$, $\phi_2(B, C)$, and $\phi_3(A, C)$ be the factors of a MN. What is

$$\sum_{A,B,C} \phi_1(A, B) \times \phi_2(B, C) \times \phi_3(A, C)?$$

a) Always less than or equal to $\phi_1(a, b) \times \phi_2(b, c) \times \phi_3(a, c)$, where a is a value of A, b is a value of B, and c is a value of C.
b) Always greater than or equal to $\phi_1(a, b) \times \phi_2(b, c) \times \phi_3(a, c)$, where a is a value of A, b is a value of B, and c is a value of C.
c) Always greater than or equal to 0
d) Always greater than or equal to 1
e) Always equal to the partition function, Z
f) Always equal to 1

** More than 1 option might be valid

Markov Networks

Probabilistic Graphical Models

Jerónimo Hernández-González