

Graphical Models

Belief Propagation Algorithm

Clément Bonnet, Charbel-Raphaël Segerie

March 31st, 2021

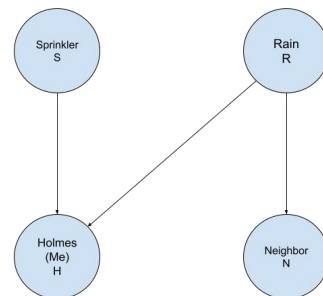
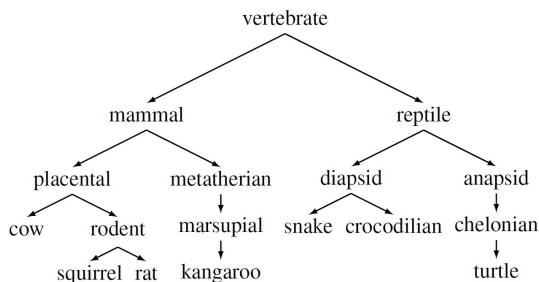
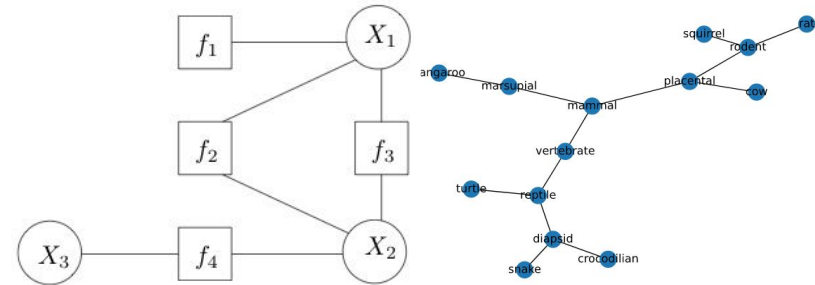
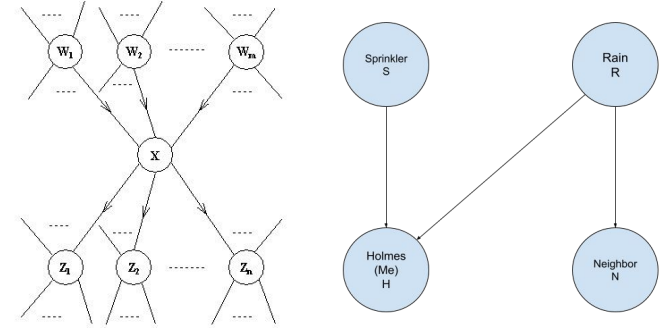


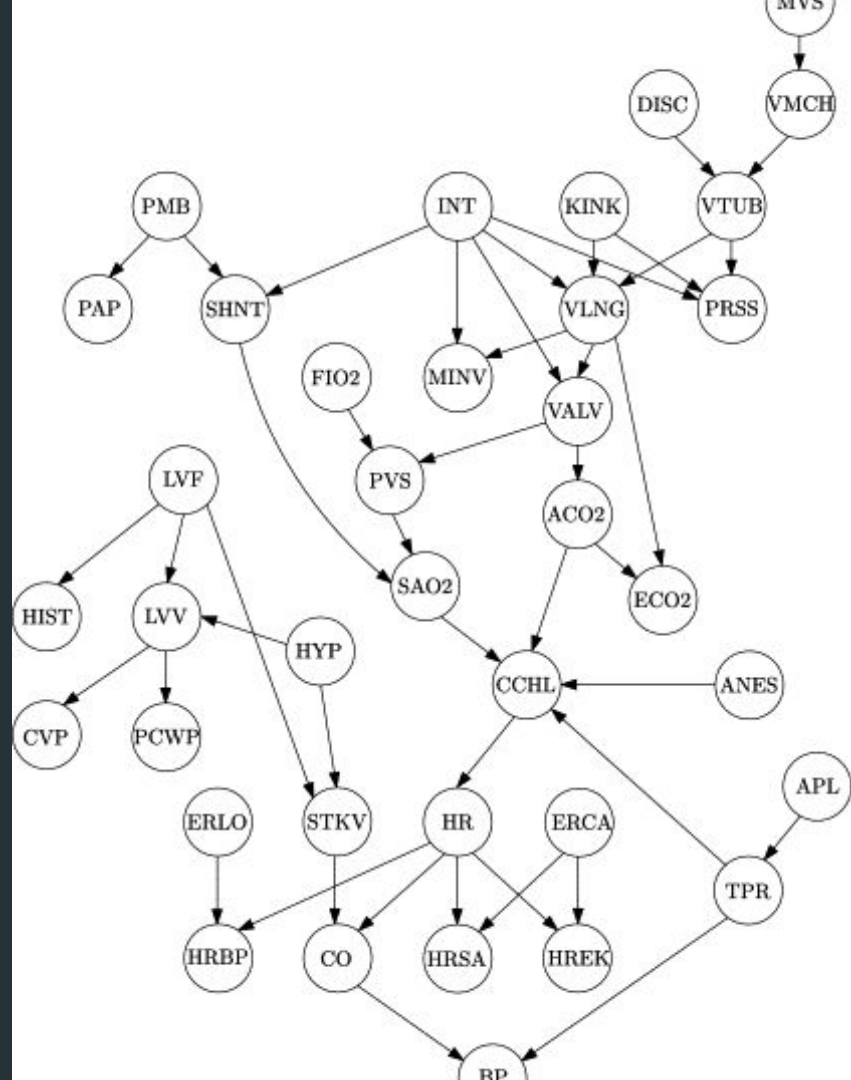
Table of Contents

1. Belief Propagation on Bayesian Networks
 - a. Theory
 - b. Case Study

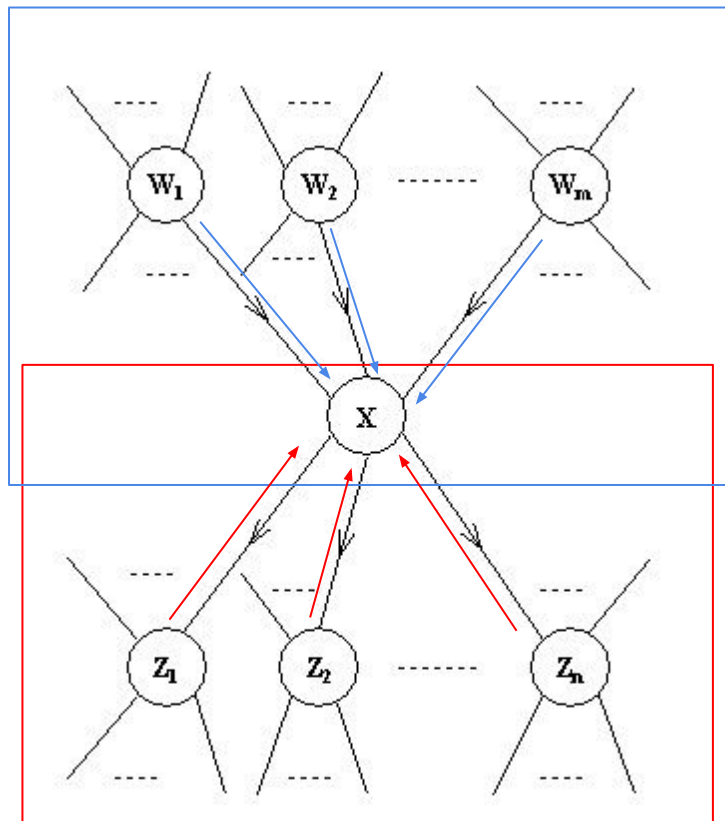
2. Belief Propagation on Factor graphs
 - a. Theory
 - b. Taxonomy Induction



Bayesian Networks



Bayesian Networks: Propagation Rules



Acyclic graph, but messages flow in both direction.

message Child -> Parent

Likelihood $\lambda(x) = \mathbb{P}(Z_1, \dots, Z_n | X=x) = \prod_{j=1}^n \lambda_{Z_j}(x)$

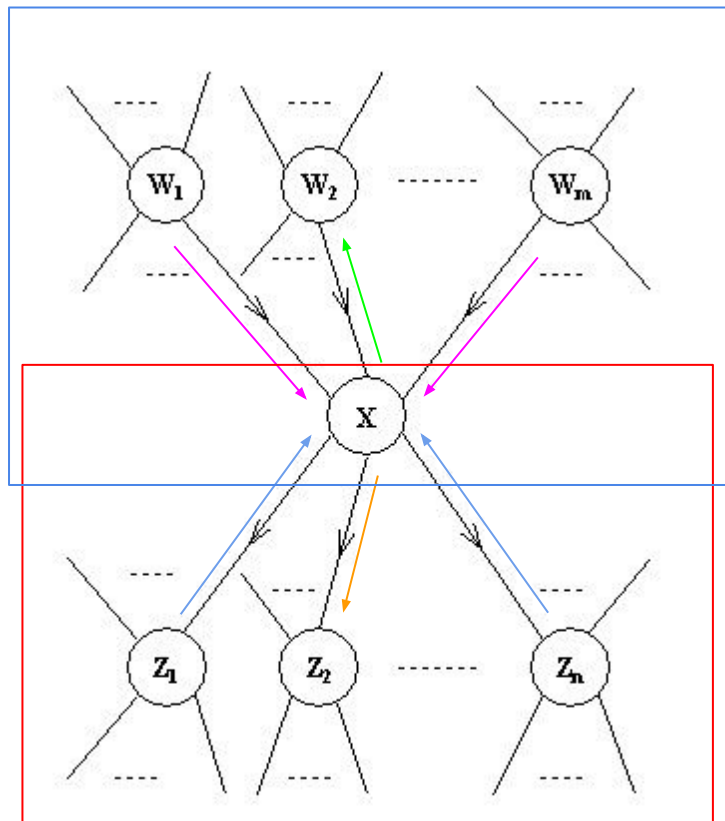
message Parent-> Child

Prior $\pi(x) = \sum_{\mathbf{w} \in \{0,1\}^m} \mathbb{P}(X=x | \mathbf{w}) \prod_{k=1}^m \pi_X(w_k)$

Belief $BEL(x) \equiv \alpha \lambda(x) \pi(x)$

$$= \alpha \left[\prod_{j=1}^n \lambda_{Z_j}(x) \right] \left[\sum_{\mathbf{w}} P(x | \mathbf{w}) \prod_{k=1}^m \pi_X(w_k) \right]$$

Bayesian Networks: Updates



Likelihood:

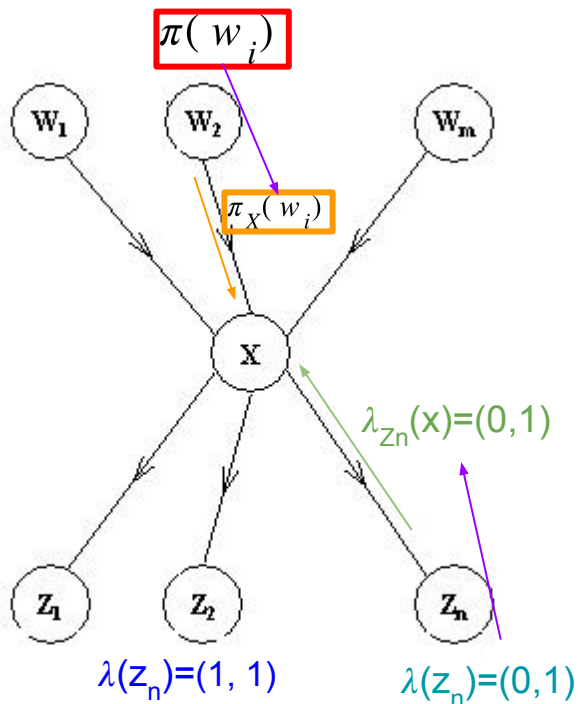
Model of the world specified by the user

$$\lambda_X(w_i) \propto \sum_x \lambda(x) \sum_{\substack{\mathbf{w}' \in \{0,1\}^m \\ w'_i = w_i}} \mathbb{P}(X = x | \mathbf{w}') \prod_{k \neq i} \pi_X(w'_k)$$

Prior:

$$\pi_{Z_j}(x) \propto \pi(x) \prod_{k \neq j} \lambda_{Z_k}(x)$$

Bayesian Networks: Boundary Conditions



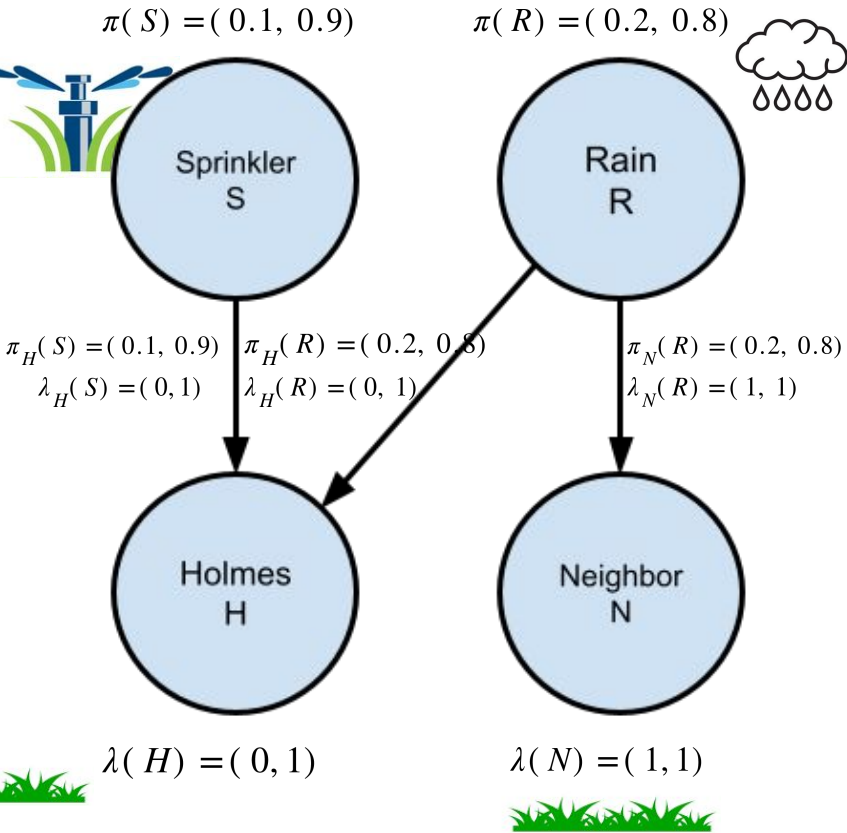
Example of binary states

Root nodes: If W is a node with no parents, we **set** the value of the **prior message** $\pi_X(w)$ equal to the **prior probabilities**.

Anticipatory nodes: If X is a childless node that has **not been instantiated**, we set the $\lambda(z)$ value as a vector of all 1's

Evidence nodes: If **evidence $X=i$ is obtained**, we set the $\lambda(z)$ value to $(0, \dots, 0, 1, 0, \dots, 0)$ with 1 at the i_{th} position, and we **set** the **likelihood message** equal to the evidence node.

Bayesian Networks: Case Study → The Wet Grass Example: Priors



Model of the world

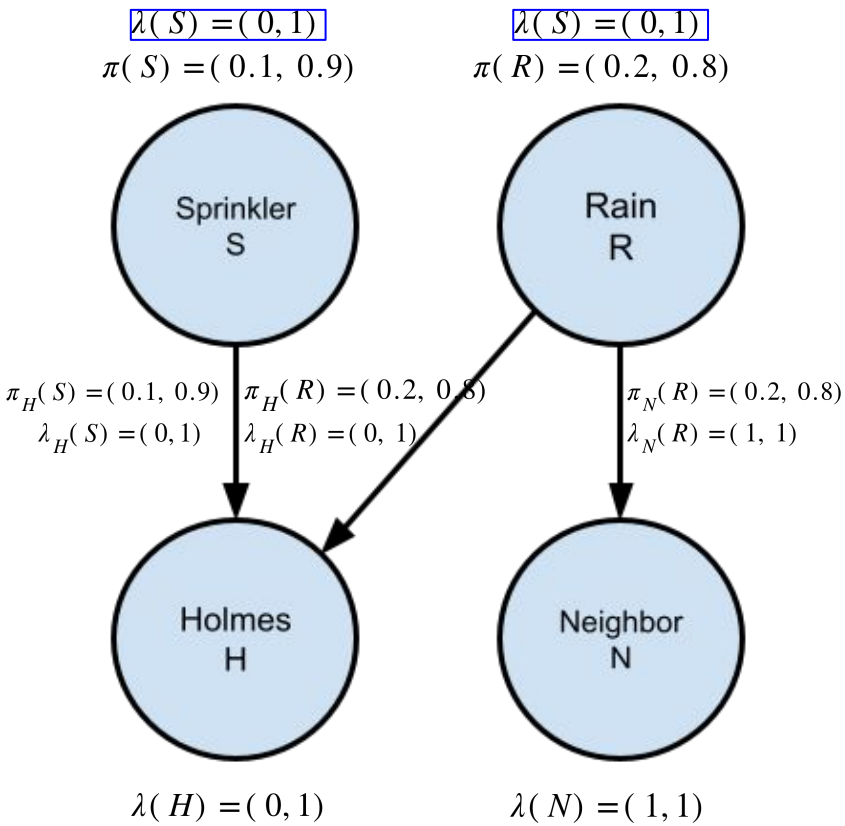
$\mathbb{P}(N R)$	$N = 0$	$N = 1$
$R = 0$	0.8	0.2
$R = 1$	0	1

$\mathbb{P}(H R, S)$	$H = 0$	$H = 1$
$R = 1, S = 1$	0	1
$R = 0, S = 1$	0.1	0.9
$R = 1, S = 0$	0	1
$R = 0, S = 0$	1	0

X	$BEL(x)$	$\pi(x)$	$\lambda(x)$
R	-	(0.2, 0.8)	-
S	-	(0.1, 0.9)	-
N	-	-	(1, 1)
H	-	-	(0, 1)

Boundary Conditions

Bayesian Networks: Case Study → The Wet Grass Example: Likelihood



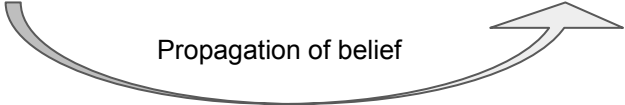
Likelihood

$$\lambda(x) = \mathbb{P}(Z_1, \dots, Z_n | X=x) = \prod_{j=1}^n \lambda_{Z_j}(x)$$

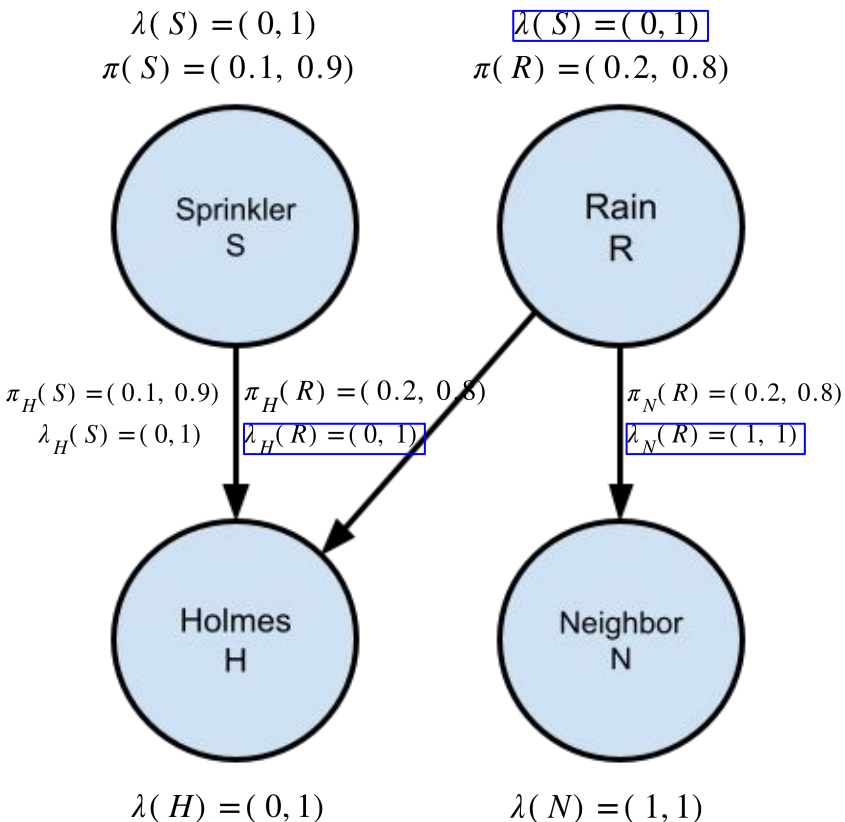
message Child → Parent

X	BEL(x)	$\pi(x)$	$\lambda(x)$
R	-	(0.2, 0.8)	-
S	-	(0.1, 0.9)	-
N	-	-	(1, 1)
H	-	-	(0, 1)

X	BEL(x)	$\pi(x)$	$\lambda(x)$
R	(0, 0.8)	(0.2, 0.8)	(0, 1)
S	(0, 0.9)	(0.1, 0.9)	(0, 1)
N	(0.16, 0.84)	(0.16, 0.84)	(1, 1)
H	(0, 0.962)	(0.038, 0.962)	(0, 1)



Bayesian Networks: Case Study → The Wet Grass Example: Likelihood



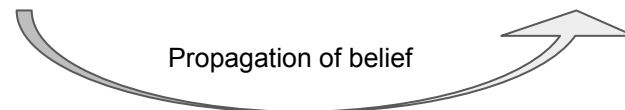
message Child → Parent

Likelihood

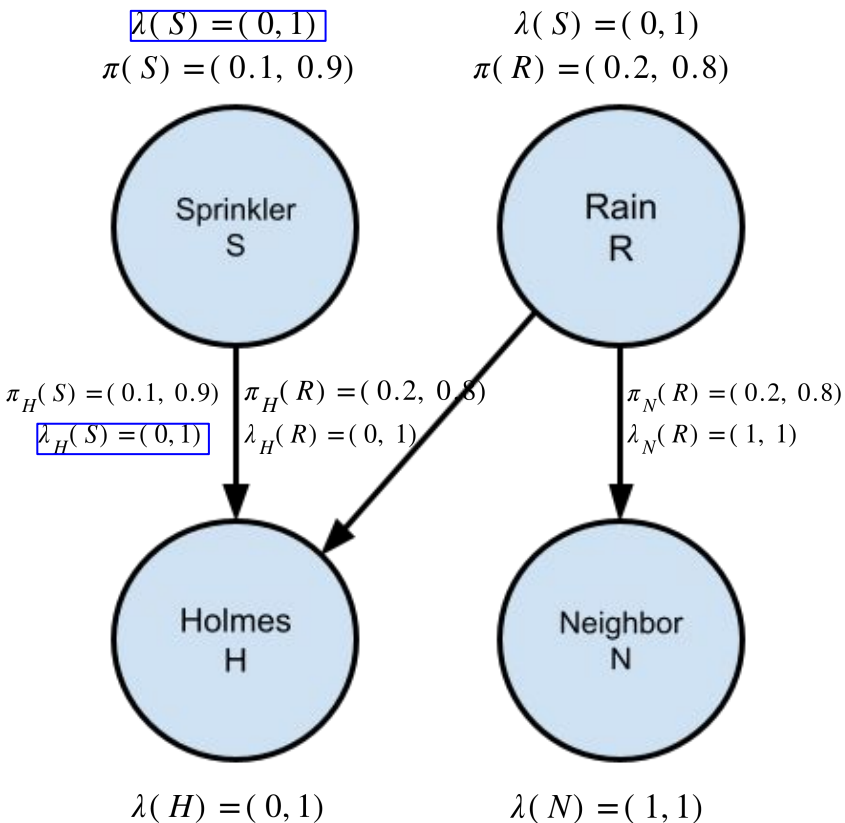
$$\lambda(x) = \mathbb{P}(Z_1, \dots, Z_n | X=x) = \prod_{j=1}^n \lambda_{Z_j}(x)$$

X	BEL(x)	$\pi(x)$	$\lambda(x)$
R	-	(0.2, 0.8)	-
S	-	(0.1, 0.9)	-
N	-	-	(1, 1)
H	-	-	(0, 1)

X	BEL(x)	$\pi(x)$	$\lambda(x)$
R	(0, 0.8)	(0.2, 0.8)	(0, 1)
S	(0, 0.9)	(0.1, 0.9)	(0, 1)
N	(0.16, 0.84)	(0.16, 0.84)	(1, 1)
H	(0, 0.962)	(0.038, 0.962)	(0, 1)



Bayesian Networks: Case Study → The Wet Grass Example: Likelihood



Likelihood

$$\lambda(x) = \mathbb{P}(Z_1, \dots, Z_n | X=x) = \prod_{j=1}^n \lambda_{Z_j}(x)$$

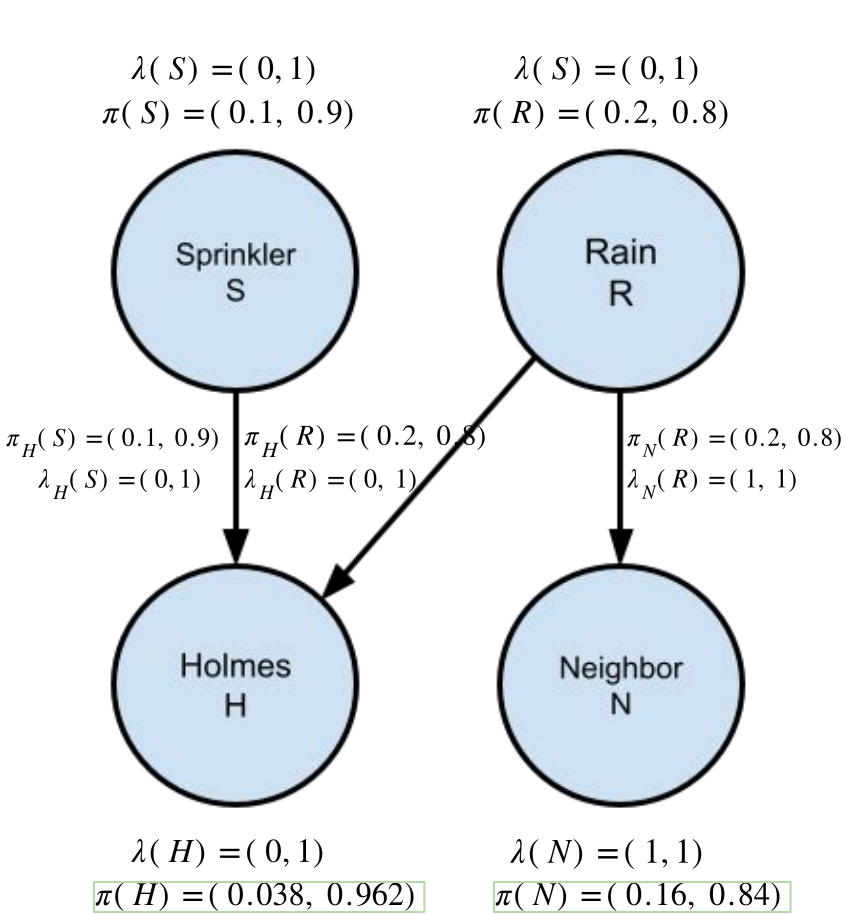
message Child → Parent

X	BEL(x)	$\pi(x)$	$\lambda(x)$
R	-	(0.2, 0.8)	-
S	-	(0.1, 0.9)	-
N	-	-	(1, 1)
H	-	-	(0, 1)

X	BEL(x)	$\pi(x)$	$\lambda(x)$
R	(0, 0.8)	(0.2, 0.8)	(0, 1)
S	(0, 0.9)	(0.1, 0.9)	(0, 1)
N	(0.16, 0.84)	(0.16, 0.84)	(1, 1)
H	(0, 0.962)	(0.038, 0.962)	(0, 1)

Propagation of belief

Bayesian Networks: Case Study: The Wet Grass Example: Priors messages



Model of the world

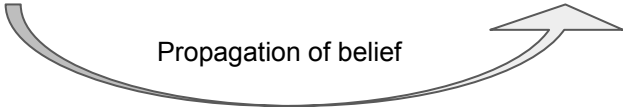
$\mathbb{P}(N R)$	$N = 0$	$N = 1$
$R = 0$	0.8	0.2
$R = 1$	0	1

$\mathbb{P}(H R, S)$	$H = 0$	$H = 1$
$R = 1, S = 1$	0	1
$R = 0, S = 1$	0.1	0.9
$R = 1, S = 0$	0	1
$R = 0, S = 0$	1	0

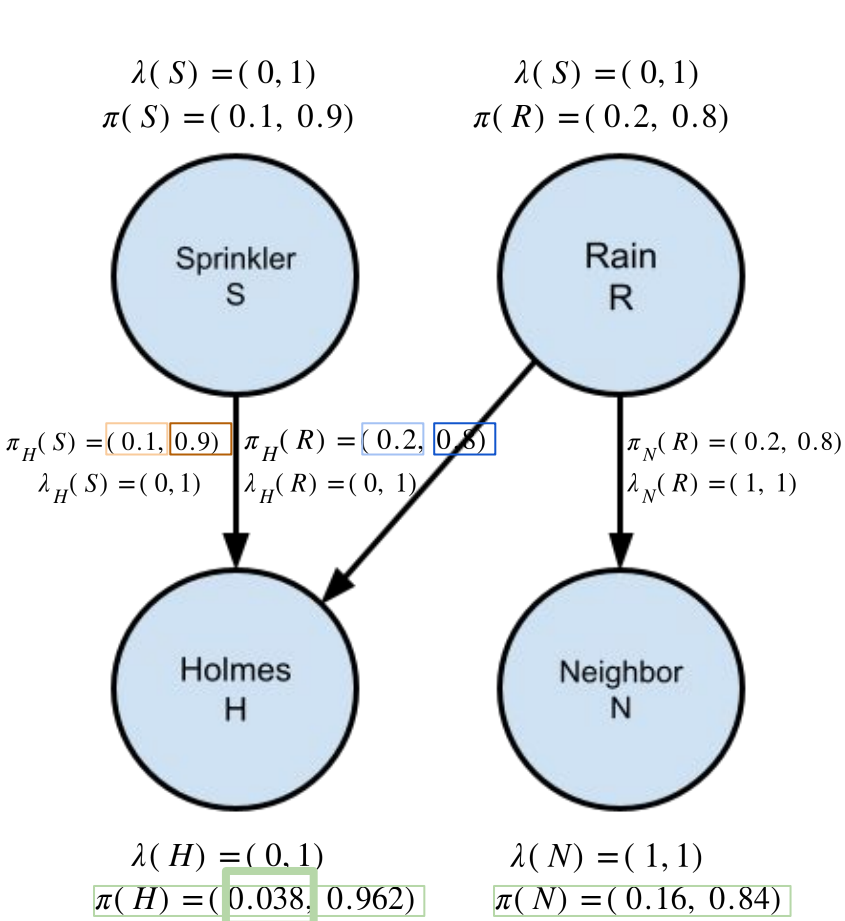
Prior
$$\pi(x) = \sum_{\mathbf{w} \in \{0,1\}^m} \mathbb{P}(X=x|\mathbf{w}) \prod_{k=1}^m \pi_X(w_k)$$

X	BEL(x)	$\pi(x)$	$\lambda(x)$
R	-	(0.2, 0.8)	-
S	-	(0.1, 0.9)	-
N	-	-	(1, 1)
H	-	-	(0, 1)

X	BEL(x)	$\pi(x)$	$\lambda(x)$
R	(0, 0.8)	(0.2, 0.8)	(0, 1)
S	(0, 0.9)	(0.1, 0.9)	(0, 1)
N	(0.16, 0.84)	(0.16, 0.84)	(1, 1)
H	(0, 0.962)	(0.038, 0.962)	(0, 1)



Bayesian Networks: Case Study: The Wet Grass Example: Priors messages



$\mathbb{P}(N R)$	$N = 0$	$N = 1$
$R = 0$	0.8	0.2
$R = 1$	0	1

Model of the world

$\mathbb{P}(H R, S)$	$H = 0$	$H = 1$
$R = 1, S = 1$	0	1
$R = 0, S = 1$	0.1	0.9
$R = 1, S = 0$	0	1
$R = 0, S = 0$	1	0

$$\text{Prior } \pi(x) = \sum_{\mathbf{w} \in \{0,1\}^m} \mathbb{P}(X=x|\mathbf{w}) \prod_{k=1}^m \pi_X(w_k)$$

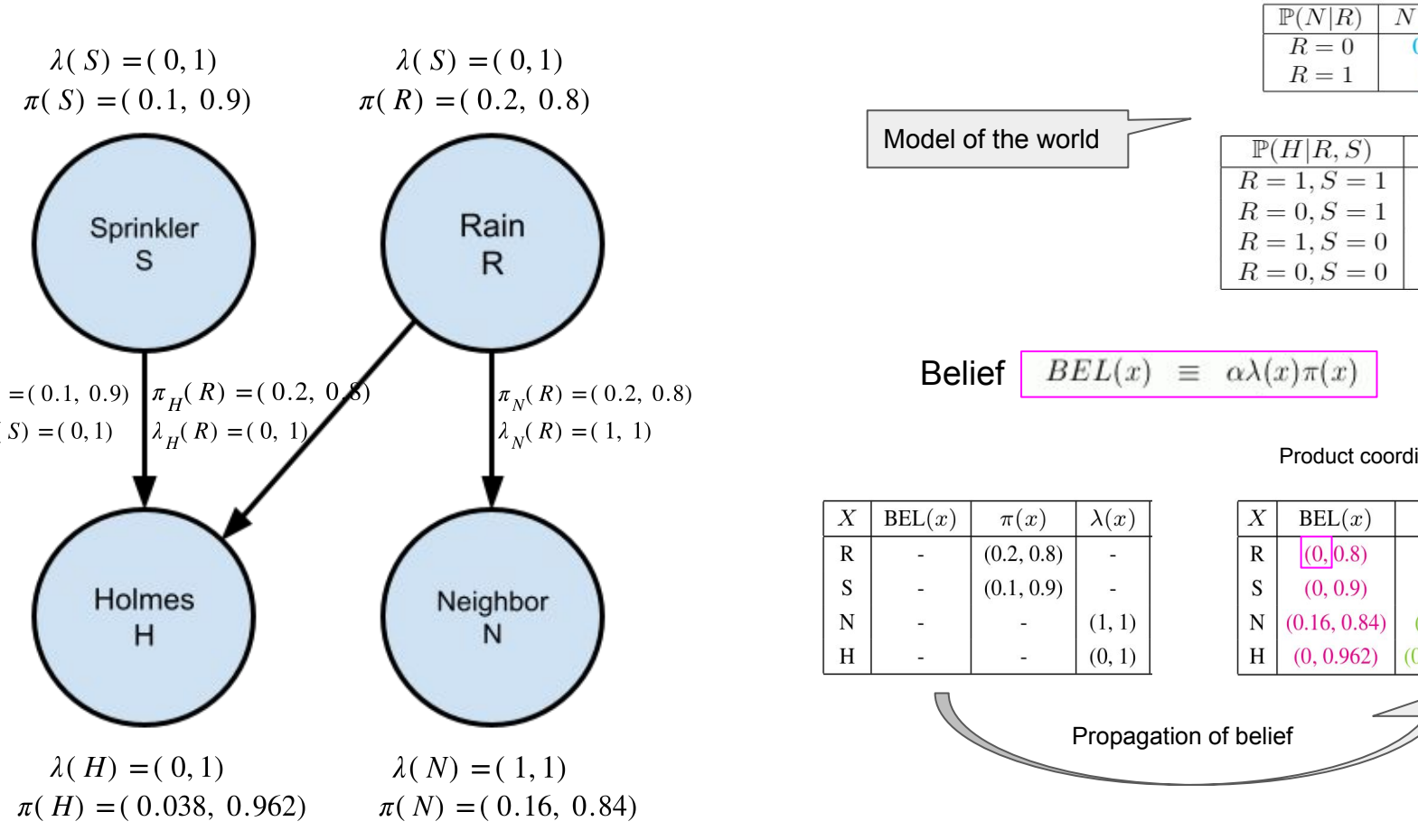
$$\begin{aligned}
 \pi(H=0) &= \mathbb{P}(H=0|R=0, S=0) \pi_H(R=0) \pi_H(S=0) \\
 &+ \mathbb{P}(H=0|R=0, S=1) \pi_H(R=0) \pi_H(S=1) \\
 &+ \mathbb{P}(H=0|R=1, S=0) \pi_H(R=1) \pi_H(S=0) \\
 &+ \mathbb{P}(H=0|R=1, S=1) \pi_H(R=1) \pi_H(S=1) \\
 &= 1 \times (0.2 \times 0.1) + 0.1 \times (0.2 \times 0.9) \\
 &+ 0 \times (0.8 \times 0.1) + 0 \times (0.8 \times 0.9) \\
 &= 0.02 + 0.018 = 0.038
 \end{aligned}$$

Graphical Model Diagram:

- Nodes:** Sprinkler (S), Rain (R), Holmes (H), Neighbor (N).
- Parents and Children:**
 - Sprinkler (S) is a parent of Holmes (H).
 - Rain (R) is a parent of Holmes (H) and Neighbor (N).
- Distributions:**
 - Marginal Distributions:**
 - $\lambda(S) = (0, 1)$
 - $\pi(S) = (0.1, 0.9)$
 - $\lambda(R) = (0, 1)$
 - $\pi(R) = (0.2, 0.8)$
 - $\lambda(H) = (0, 1)$
 - $\pi(H) = (0.038, 0.962)$
 - $\lambda(N) = (1, 1)$
 - $\pi(N) = (0.16, 0.84)$
 - Conditional Distributions:**
 - $\pi_H(R) = (0.2, 0.8)$
 - $\lambda_H(R) = (0, 1)$
 - $\pi_N(R) = (0.2, 0.8)$
 - $\lambda_N(R) = (1, 1)$

Table: $\mathbb{P}(H, R, S)$

R	S	H	$\mathbb{P}(H, R, S)$
1	1	0	0.000
1	1	1	0.000
1	0	0	0.000
1	0	1	0.000
0	1	0	0.000
0	1	1	0.000
0	0	0	0.000
0	0	1	0.000



Graphical Model:

- Nodes:** Sprinkler S , Rain R , Holmes H , Neighbor N .
- Edges:** $S \rightarrow H$, $R \rightarrow H$, $R \rightarrow N$.
- Node Distributions:**
 - $\lambda(S) = (0, 1)$, $\pi(S) = (0.1, 0.9)$
 - $\lambda(R) = (0, 1)$, $\pi(R) = (0.2, 0.8)$
 - $\lambda(H) = (0, 1)$, $\pi(H) = (0.038, 0.962)$
 - $\lambda(N) = (1, 1)$, $\pi(N) = (0.16, 0.84)$
- Edge Labels:**
 - $S \rightarrow H$: $\pi_H(R) = (0.2, 0.8)$, $\lambda_H(R) = (0, 1)$
 - $R \rightarrow H$: $\pi_N(R) = (0.2, 0.8)$, $\lambda_N(R) = (1, 1)$

Model of the world

$\mathbb{P}(H R, S)$	
$R = 1, S = 1$	0.962
$R = 0, S = 1$	0.9
$R = 1, S = 0$	0.84
$R = 0, S = 0$	0.038

Belief $BEL(x) \equiv \alpha \lambda(x) \pi(x)$

X	$BEL(x)$	$\pi(x)$	$\lambda(x)$
R	-	(0.2, 0.8)	-
S	-	(0.1, 0.9)	-
N	-	-	(1, 1)
H	-	-	(0, 1)

X	$BEL(x)$	$\pi(x)$	$\lambda(x)$
R	-	(0.2, 0.8)	-
S	-	(0.1, 0.9)	-
N	(0.16, 0.84)	-	(1, 1)
H	(0.038, 0.962)	-	(0, 1)

Propagation of belief

Graphical Model:

- Nodes:** Sprinkler S , Rain R , Holmes H , Neighbor N .
- Edges:** $S \rightarrow H$, $R \rightarrow H$, $R \rightarrow N$.
- Node Distributions:**
 - $\lambda(S) = (0, 1)$, $\pi(S) = (0.1, 0.9)$
 - $\lambda(R) = (0, 1)$, $\pi(R) = (0.2, 0.8)$
 - $\lambda(H) = (0, 1)$, $\pi(H) = (0.038, 0.962)$
 - $\lambda(N) = (1, 1)$, $\pi(N) = (0.16, 0.84)$
- Edge Labels:**
 - $S \rightarrow H$: $\pi_H(R) = (0.2, 0.8)$, $\lambda_H(R) = (0, 1)$
 - $R \rightarrow H$: $\pi_N(R) = (0.2, 0.8)$, $\lambda_N(R) = (1, 1)$

Model of the world

$\mathbb{P}(H R, S)$	
$R = 1, S = 1$	0.962
$R = 0, S = 1$	0.9
$R = 1, S = 0$	0.84
$R = 0, S = 0$	0.038

Belief $BEL(x) \equiv \alpha \lambda(x) \pi(x)$

X	$BEL(x)$	$\pi(x)$	$\lambda(x)$
R	-	(0.2, 0.8)	-
S	-	(0.1, 0.9)	-
N	-	-	(1, 1)
H	-	-	(0, 1)

X	$BEL(x)$	$\pi(x)$	$\lambda(x)$
R	-	(0.2, 0.8)	-
S	-	(0.1, 0.9)	-
N	(0.16, 0.84)	-	(1, 1)
H	(0.038, 0.962)	-	(0, 1)

Propagation of belief

Model of the world

Belief $BEL(x) \equiv \alpha \lambda(x) \pi(x)$

Product coordinate

X	$BEL(x)$	$\pi(x)$	$\lambda(x)$
R	-	(0.2, 0.8)	-
S	-	(0.1, 0.9)	-
N	-	-	(1, 1)
H	-	-	(0, 1)

X	$BEL(x)$
R	(0, 0.8)
S	(0, 0.9)
N	(0.16, 0.84)
H	(0, 0.962)

Propagation of belief

Graphical Model:

- Nodes:** Sprinkler S , Rain R , Holmes H , Neighbor N .
- Edges:** $S \rightarrow H$, $R \rightarrow H$, $R \rightarrow N$.
- Node Distributions:**
 - $\lambda(S) = (0, 1)$, $\pi(S) = (0.1, 0.9)$
 - $\lambda(R) = (0, 1)$, $\pi(R) = (0.2, 0.8)$
 - $\lambda(H) = (0, 1)$, $\pi(H) = (0.038, 0.962)$
 - $\lambda(N) = (1, 1)$, $\pi(N) = (0.16, 0.84)$
- Edge Labels:**
 - $S \rightarrow H$: $\pi_H(R) = (0.2, 0.8)$, $\lambda_H(R) = (0, 1)$
 - $R \rightarrow H$: $\pi_N(R) = (0.2, 0.8)$, $\lambda_N(R) = (1, 1)$

Model of the world

$\mathbb{P}(H R, S)$	
$R = 1, S = 1$	0.962
$R = 0, S = 1$	0.962
$R = 1, S = 0$	0.962
$R = 0, S = 0$	0.962

Belief $BEL(x) \equiv \alpha \lambda(x) \pi(x)$

X	$BEL(x)$	$\pi(x)$	$\lambda(x)$
R	-	(0.2, 0.8)	-
S	-	(0.1, 0.9)	-
N	-	-	(1, 1)
H	-	-	(0, 1)

Product coordinate

X	$BEL(x)$	$\pi(x)$	$\lambda(x)$
R	(0, 0.8)	(0.2, 0.8)	-
S	(0, 0.9)	(0.1, 0.9)	-
N	(0.16, 0.84)	-	(1, 1)
H	(0, 0.962)	-	(0, 1)

Propagation of belief

$\lambda(S) = (0, 1)$
 $\pi(S) = (0.1, 0.9)$

$\lambda(R) = (0, 1)$
 $\pi(R) = (0.2, 0.8)$

$\lambda(H) = (0, 1)$
 $\pi(H) = (0.038, 0.962)$

$\lambda(N) = (1, 1)$
 $\pi(N) = (0.16, 0.84)$

Model of the world

$\mathbb{P}(H R, S)$	
$R = 1, S = 1$	0.008
$R = 0, S = 1$	0.000
$R = 1, S = 0$	0.000
$R = 0, S = 0$	0.000

Belief $BEL(x) \equiv \alpha \lambda(x) \pi(x)$

X	$BEL(x)$	$\pi(x)$	$\lambda(x)$
R	-	(0.2, 0.8)	-
S	-	(0.1, 0.9)	-
N	-	-	(1, 1)
H	-	-	(0, 1)

X	$BEL(x)$
R	(0, 0.8)
S	(0, 0.9)
N	(0.16, 0.84)
H	(0, 0.962)

Propagation of belief

$\lambda(S) = (0, 1)$
 $\pi(S) = (0.1, 0.9)$

$\lambda(R) = (0, 1)$
 $\pi(R) = (0.2, 0.8)$

$\lambda(H) = (0, 1)$
 $\pi(H) = (0.038, 0.962)$

$\lambda(N) = (1, 1)$
 $\pi(N) = (0.16, 0.84)$

Joint distributions for parents of H:
 $\pi_H(R) = (0.2, 0.8)$
 $\lambda_H(R) = (0, 1)$

Joint distributions for parents of N:
 $\pi_N(R) = (0.2, 0.8)$
 $\lambda_N(R) = (1, 1)$

Model of the world

$\mathbb{P}(N R)$		N
$R = 0$		0
$R = 1$		1

$\mathbb{P}(H R, S)$	
$R = 1, S = 1$	0.962
$R = 0, S = 1$	0.038
$R = 1, S = 0$	0.038
$R = 0, S = 0$	0.962

Belief $BEL(x) \equiv \alpha \lambda(x) \pi(x)$

X	$BEL(x)$	$\pi(x)$	$\lambda(x)$
R	-	(0.2, 0.8)	-
S	-	(0.1, 0.9)	-
N	-	-	(1, 1)
H	-	-	(0, 1)

X	$BEL(x)$	
R	(0, 0.8)	
S	(0, 0.9)	
N	(0.16, 0.84)	
H	(0, 0.962)	

Product coordinate

Propagation of belief

$\lambda(S) = (0, 1)$
 $\pi(S) = (0.1, 0.9)$

$\lambda(R) = (0, 1)$
 $\pi(R) = (0.2, 0.8)$

$\lambda(H) = (0, 1)$
 $\pi(H) = (0.038, 0.962)$

$\lambda(N) = (1, 1)$
 $\pi(N) = (0.16, 0.84)$

Model of the world

$\mathbb{P}(H R, S)$	
$R = 1, S = 1$	0.008
$R = 0, S = 1$	0.000
$R = 1, S = 0$	0.000
$R = 0, S = 0$	0.000

Belief $BEL(x) \equiv \alpha \lambda(x) \pi(x)$

X	$BEL(x)$	$\pi(x)$	$\lambda(x)$
R	-	(0.2, 0.8)	-
S	-	(0.1, 0.9)	-
N	-	-	(1, 1)
H	-	-	(0, 1)

X	$BEL(x)$
R	(0, 0.8)
S	(0, 0.9)
N	(0.16, 0.84)
H	(0, 0.962)

Propagation of belief

$\lambda(S) = (0, 1)$
 $\pi(S) = (0.1, 0.9)$

$\lambda(R) = (0, 1)$
 $\pi(R) = (0.2, 0.8)$

$\lambda(H) = (0, 1)$
 $\pi(H) = (0.038, 0.962)$

$\lambda(N) = (1, 1)$
 $\pi(N) = (0.16, 0.84)$

Model of the world

$\mathbb{P}(H R, S)$	
$R = 1, S = 1$	0.008
$R = 0, S = 1$	0.000
$R = 1, S = 0$	0.000
$R = 0, S = 0$	0.000

Belief $BEL(x) \equiv \alpha \lambda(x) \pi(x)$

X	$BEL(x)$	$\pi(x)$	$\lambda(x)$
R	-	(0.2, 0.8)	-
S	-	(0.1, 0.9)	-
N	-	-	(1, 1)
H	-	-	(0, 1)

X	$BEL(x)$
R	(0, 0.8)
S	(0, 0.9)
N	(0.16, 0.84)
H	(0, 0.962)

Propagation of belief

Initial Beliefs:

- Sprinkler (S):** $\lambda(S) = (0, 1)$, $\pi(S) = (0.1, 0.9)$
- Rain (R):** $\lambda(R) = (0, 1)$, $\pi(R) = (0.2, 0.8)$
- Holmes (H):** $\lambda(H) = (0, 1)$, $\pi(H) = (0.038, 0.962)$
- Neighbor (N):** $\lambda(N) = (1, 1)$, $\pi(N) = (0.16, 0.84)$

Model of the world:

$\mathbb{P}(H R,S)$	$H = 0$	$H = 1$
$R = 1, S = 1$	0	1
$R = 0, S = 1$	0.1	0.9
$R = 1, S = 0$	0	1
$R = 0, S = 0$	1	0

Belief Propagation:

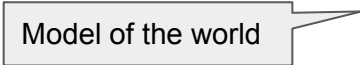
Belief: $BEL(x) \equiv \alpha \lambda(x) \pi(x)$

X	$BEL(x)$	$\pi(x)$	$\lambda(x)$
R	-	(0.2, 0.8)	-
S	-	(0.1, 0.9)	-
N	-	-	(1, 1)
H	-	-	(0, 1)

Propagated Beliefs:


X	$BEL(x)$	$\pi(x)$	$\lambda(x)$
R	(0, 0.8)	(0.2, 0.8)	(0, 1)
S	(0, 0.9)	(0.1, 0.9)	(0, 1)
N	(0.16, 0.84)	(0.16, 0.84)	(1, 1)
H	(0, 0.962)	(0.038, 0.962)	(0, 1)

Propagation of belief



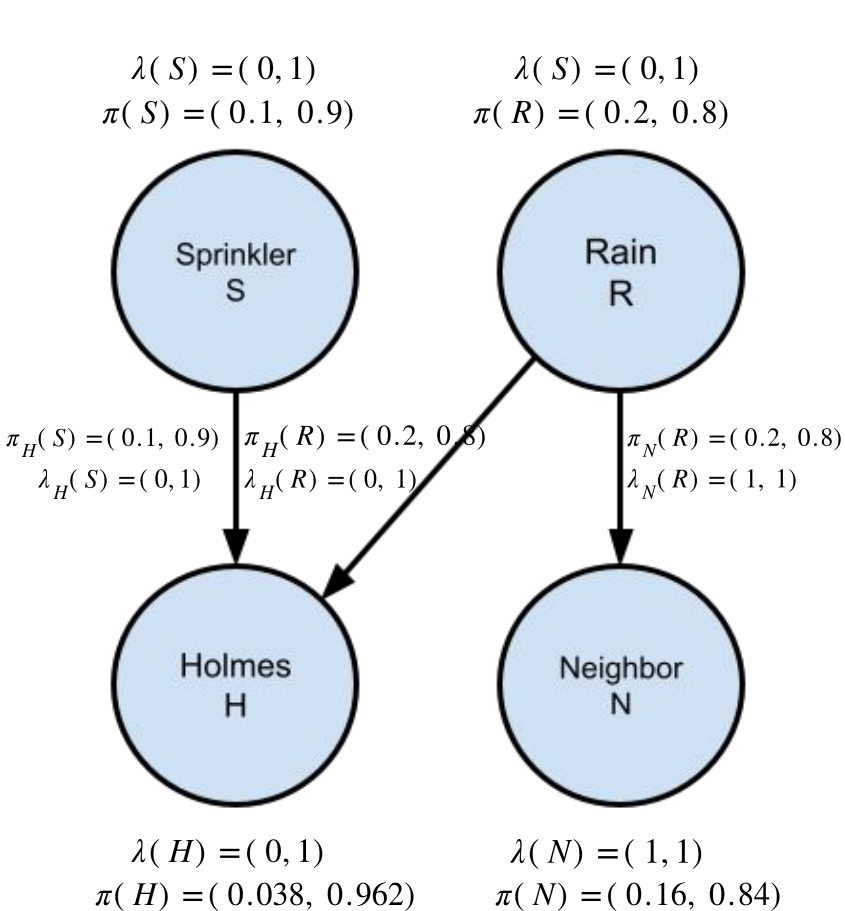
$\mathbb{P}(H R, S)$	$H = 0$	$H = 1$
$R = 1, S = 1$	0	1
$R = 0, S = 1$	0.1	0.9
$R = 1, S = 0$	0	1
$R = 0, S = 0$	1	0

X	$\text{BEL}(x)$	$\pi(x)$	$\lambda(x)$
R	-	(0.2, 0.8)	-
S	-	(0.1, 0.9)	-
N	-	-	(1, 1)
H	-	-	(0, 1)



Propagation of belief

Bayesian Networks: Case Study: The Wet Grass Example: Beliefs



Model of the world

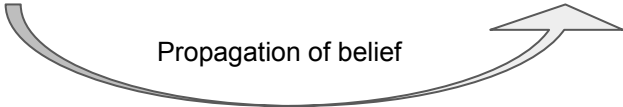
$\mathbb{P}(N R)$	$N = 0$	$N = 1$
$R = 0$	0.8	0.2
$R = 1$	0	1

$\mathbb{P}(H R, S)$	$H = 0$	$H = 1$
$R = 1, S = 1$	0	1
$R = 0, S = 1$	0.1	0.9
$R = 1, S = 0$	0	1
$R = 0, S = 0$	1	0

Belief $BEL(x) \equiv \alpha \lambda(x) \pi(x)$

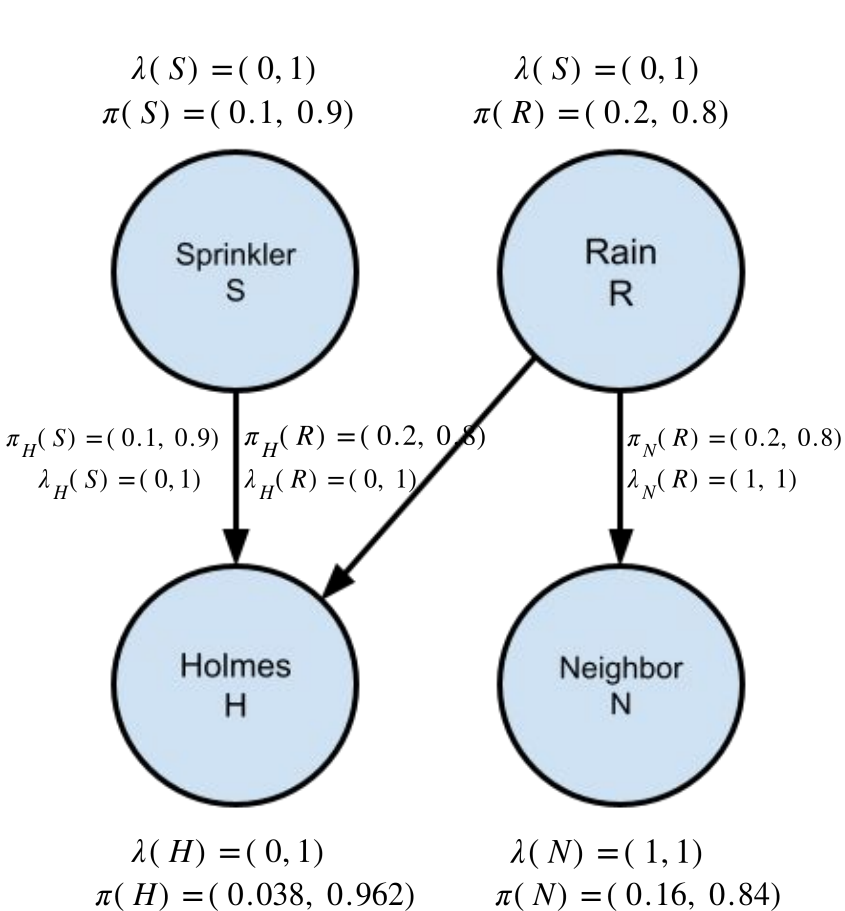
X	$BEL(x)$	$\pi(x)$	$\lambda(x)$
R	-	(0.2, 0.8)	-
S	-	(0.1, 0.9)	-
N	-	-	(1, 1)
H	-	-	(0, 1)

X	$BEL(x)$	$\pi(x)$	$\lambda(x)$
R	(0, 0.8)	(0.2, 0.8)	(0, 1)
S	(0, 0.9)	(0.1, 0.9)	(0, 1)
N	(0.16, 0.84)	(0.16, 0.84)	(1, 1)
H	(0, 0.962)	(0.038, 0.962)	(0, 1)



Propagation of belief

Bayesian Networks: Case Study: The Wet Grass Example: Beliefs



Model of the world

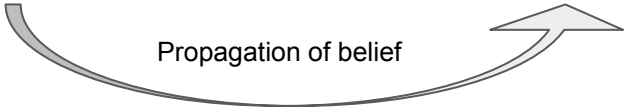
$\mathbb{P}(N R)$	$N = 0$	$N = 1$
$R = 0$	0.8	0.2
$R = 1$	0	1

$\mathbb{P}(H R, S)$	$H = 0$	$H = 1$
$R = 1, S = 1$	0	1
$R = 0, S = 1$	0.1	0.9
$R = 1, S = 0$	0	1
$R = 0, S = 0$	1	0

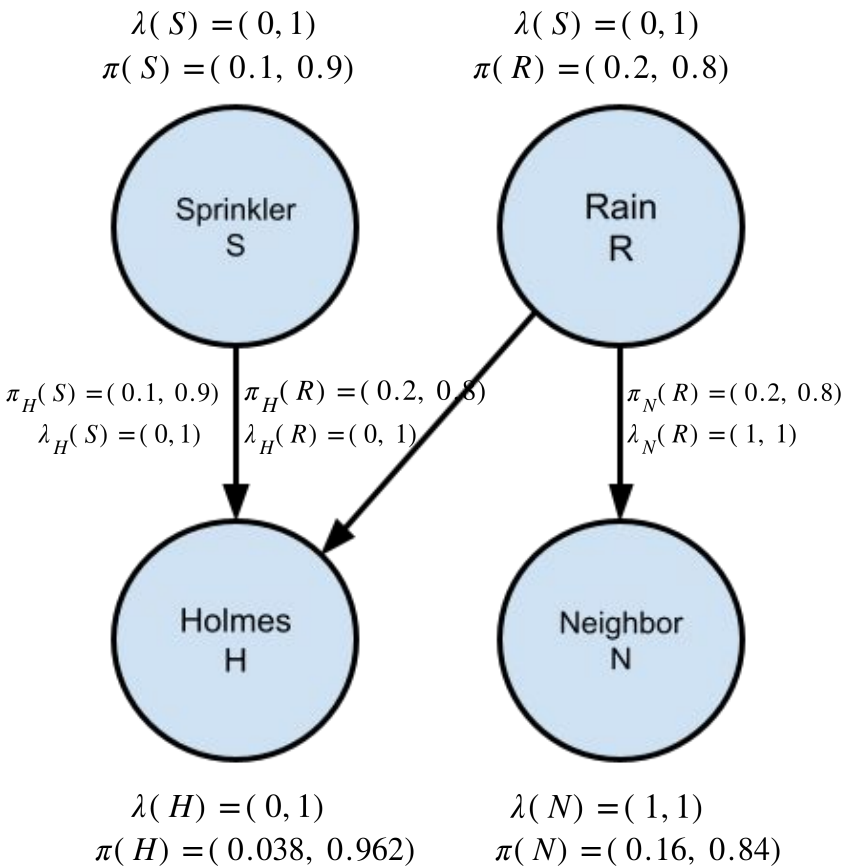
Belief $BEL(x) \equiv \alpha \lambda(x) \pi(x)$

X	BEL(x)	$\pi(x)$	$\lambda(x)$
R	-	(0.2, 0.8)	-
S	-	(0.1, 0.9)	-
N	-	-	(1, 1)
H	-	-	(0, 1)

X	BEL(x)	$\pi(x)$	$\lambda(x)$
R	(0, 0.8)	(0.2, 0.8)	(0, 1)
S	(0, 0.9)	(0.1, 0.9)	(0, 1)
N	(0.16, 0.84)	(0.16, 0.84)	(1, 1)
H	(0, 0.962)	(0.038, 0.962)	(0, 1)



Bayesian Networks: Case Study: The Wet Grass Example: Beliefs



$\mathbb{P}(N R)$	$N = 0$	$N = 1$
$R = 0$	0.8	0.2
$R = 1$	0	1

Model of the world

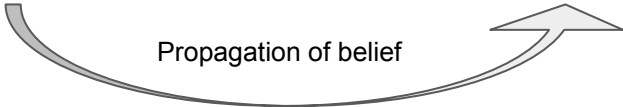
$\mathbb{P}(H R, S)$	$H = 0$	$H = 1$
$R = 1, S = 1$	0	1
$R = 0, S = 1$	0.1	0.9
$R = 1, S = 0$	0	1
$R = 0, S = 0$	1	0

Belief $BEL(x) \equiv \alpha \lambda(x) \pi(x)$

etc...

X	$BEL(x)$	$\pi(x)$	$\lambda(x)$
R	-	(0.2, 0.8)	-
S	-	(0.1, 0.9)	-
N	-	-	(1, 1)
H	-	-	(0, 1)

X	$BEL(x)$	$\pi(x)$	$\lambda(x)$
R	(0, 0.8)	(0.2, 0.8)	(0, 1)
S	(0, 0.9)	(0.1, 0.9)	(0, 1)
N	(0.16, 0.84)	(0.16, 0.84)	(1, 1)
H	(0, 0.962)	(0.038, 0.962)	(0, 1)



Propagation of belief

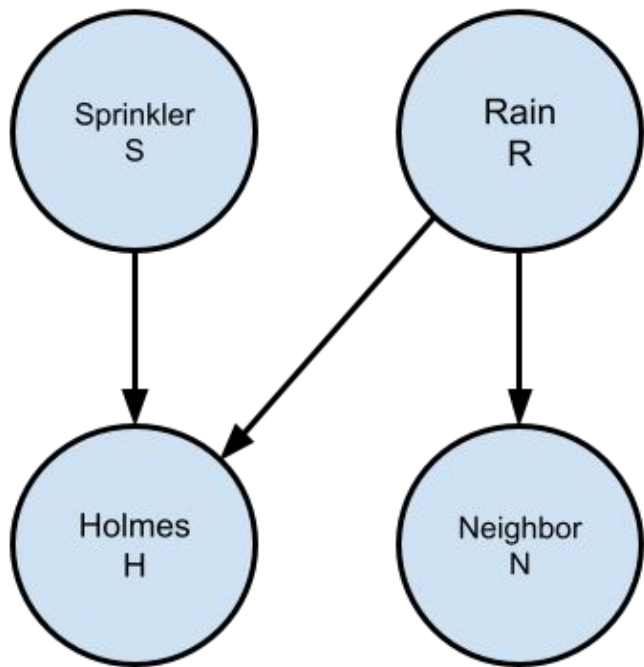
Bayesian Networks: Case Study: The Wet Grass Example: Results

Holmes sees only his grass

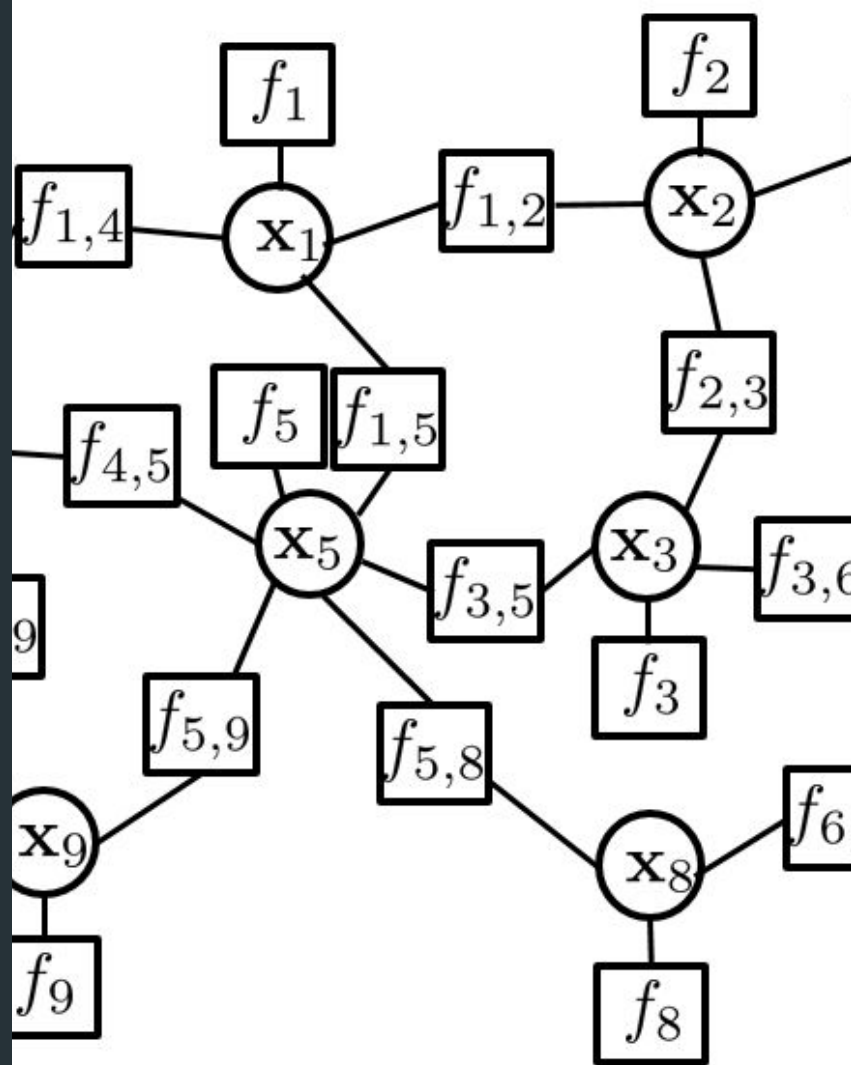
$$\left\{ \begin{array}{l} \text{Watson's grass is wet: } 0.865 \\ \text{Holmes' grass is wet: } 1 \\ \text{It has rained: } \underline{0.832} \\ \text{The sprinkler was left on: } 0.917 \end{array} \right.$$

If Holmes also sees his neighbor's wet lawn

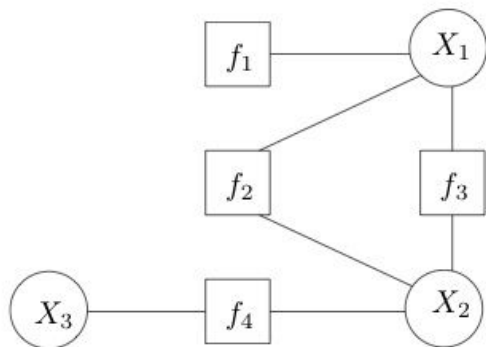
$$\left\{ \begin{array}{l} \text{Watson's grass is wet: } 1 \\ \text{Holmes' grass is wet: } 1 \\ \text{It has rained: } \underline{0.961} \\ \text{The sprinkler was left on: } 0.904 \end{array} \right.$$



Factor Graphs



Factor Graphs: Theory



Factor Graph = Bipartite Graph

□ Factor node

○ Variable node

Total likelihood:

$$p(\mathbf{x}) = \prod_{f \in F} \mathbf{f}_f(\mathbf{x}_f) = f_1(X_1) f_2(X_1, X_2) f_3(X_1, X_2) f_4(X_2, X_3)$$

Propagation:

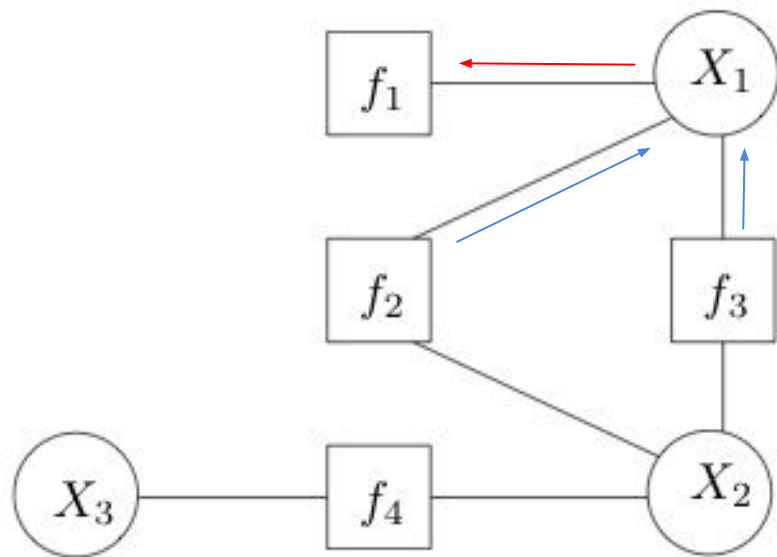
$$\mu_{v \rightarrow f}(x_v) = \prod_{f^* \in N(v) \setminus \{f\}} \mu_{f^* \rightarrow v}(x_v)$$

$$\mu_{f \rightarrow v}(x_v) = \sum_{\mathbf{x}'_f : x'_v = x_v} \mathbf{f}_f(\mathbf{x}'_f) \prod_{v^* \in N(f) \setminus \{v\}} \mu_{v^* \rightarrow f}(x'_{v^*})$$

Marginals:

$$p_{X_v}(x_v) \propto \prod_{f \in N(v)} \mu_{f \rightarrow v}(x_v)$$

Factor Graphs: Example of Propagation: Variable-to-Factor Node



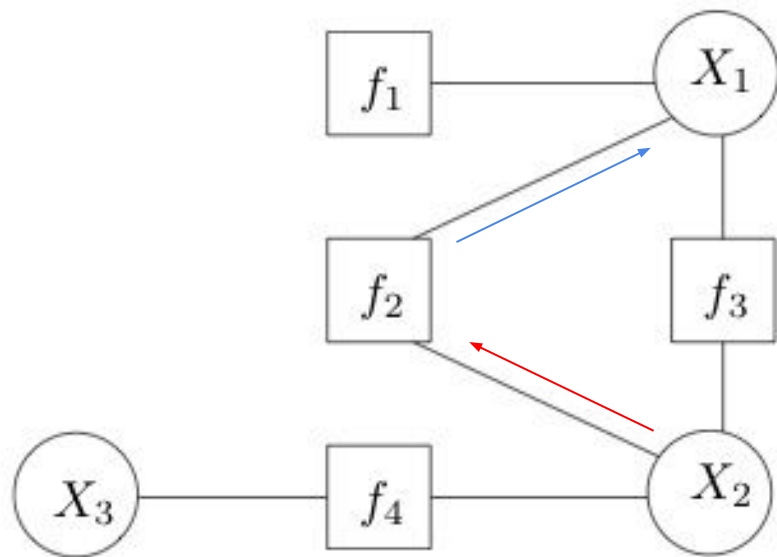
$$\overrightarrow{\mu_{v \rightarrow f}(x_v)} = \prod_{f^* \in N(v) \setminus \{f\}} \overrightarrow{\mu_{f^* \rightarrow v}(x_v)}$$

Messages from variable X to factor nodes f are vectors of size $\text{NbStates}(X)$

Messages from factor f to variable X are also vectors of size $\text{NbStates}(X)$

We just multiply coordinate by coordinate the two incoming messages

Factor Graphs: Example of Propagation: Factor-to-Variable Node



$$\mu_{f \rightarrow v}(x_v) = \sum_{\mathbf{x}'_f: \mathbf{x}'_v = x_v} f_f(\mathbf{x}'_f) \prod_{v^* \in N(f) \setminus \{v\}} \mu_{v^* \rightarrow f}(x'_{v^*})$$

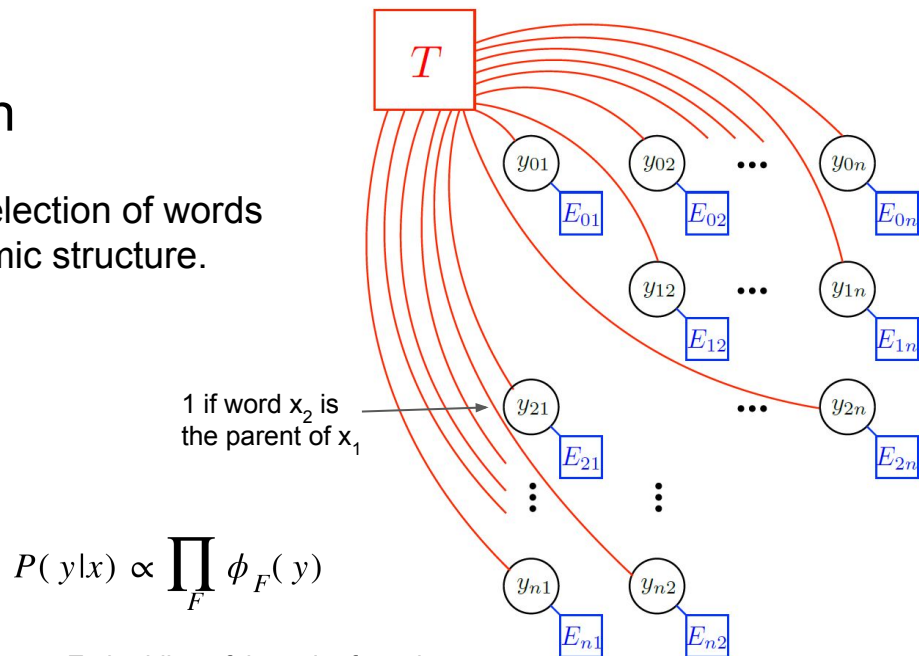
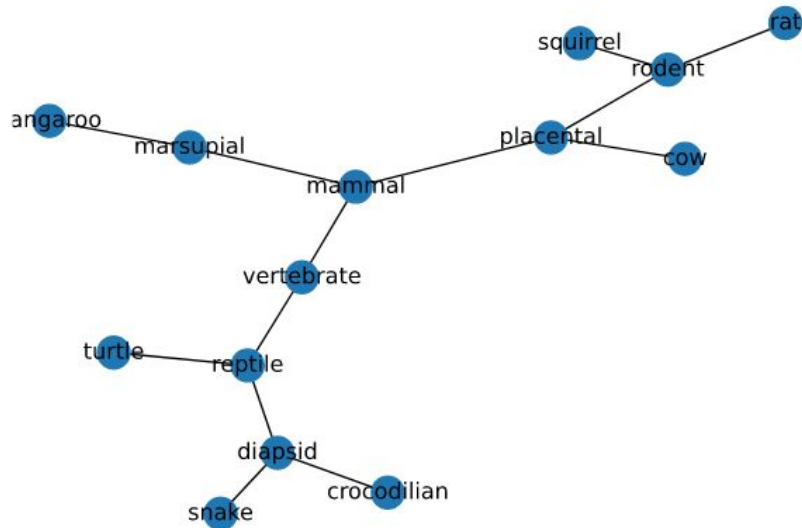
Messages from variable X to factor nodes f are vectors of size $\text{NbStates}(X)$

Messages from factor f to variable X are also vectors of size $\text{NbStates}(X)$

We just multiply coordinate by coordinate the **incoming messages**

Factor Graphs: Taxonomy Induction

From a corpus of abstracts on Wikipedia, and a selection of words (x_1, \dots, x_n), the objective is to discover the taxonomic structure.



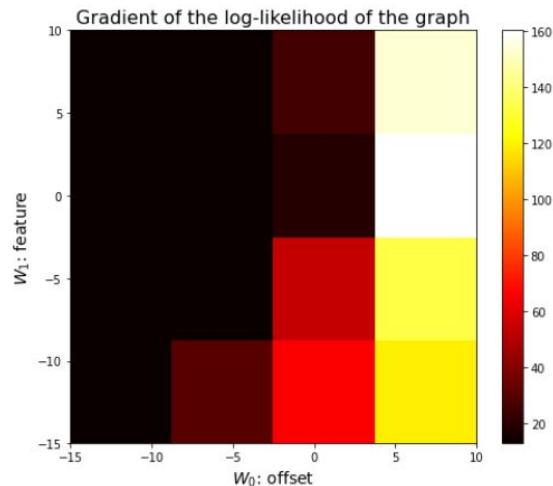
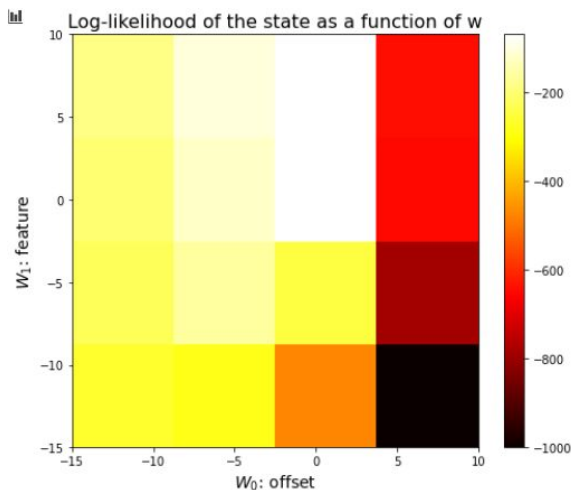
$$P(y|x) \propto \prod_F \phi_F(y)$$

Embedding of the pair of words

$$\phi_{E_{i,j}}(y_{i,j}) = \begin{cases} \exp(\mathcal{U}(x_i, x_j) \cdot w) & \text{if } y_{i,j} = 1 \\ \exp(0) = 1 & \text{if } y_{i,j} = 0 \end{cases}$$

$$\phi_T(y_{i,j}) = \begin{cases} 1 & \text{if } y \text{ forms a legal taxonomy tree} \\ 0 & \text{otherwise} \end{cases}$$

Factor Graphs: Taxonomy Induction \rightarrow Training



- Trained on a taxonomic structure by maximizing its likelihood
- Gradient Descent computed by finite difference
- Trained on a few iterations upon convergence
- Yet, too many edges are drawn

Conclusion

Implemented belief propagation on two types of graphs:

- Bayesian networks → case study with the wet grass example
- Factor graphs → taxonomy induction