지식표현과 추론

확률 그래프 모델과 지식 표현의 문제

이건명

충북대학교 소프트웨어학과

인공지능 : 튜링 테스트에서 딥러닝까지

학습 내용

- 확률 그래프 모델의 개념에 대해서 알아본다.
- 베이지안 망, 마르코프 랜덤 필드, 조건부 랜덤 필드, 로그-선형 모델의 형태에 대해서 알아본다.
- 심볼 그라운딩 문제와 프레임 문제에 대해서 알아본다.
- 상식의 표현에 대해서 알아본다.

1. 확률 그래프 모델

❖ **확률 그래프 모델(probabilistic graphical model)**

- **확률 이론**과 **그래프 이론**을 결합하여 **확률분포**(probability distribution)를 **표현**하고,
  관심있는 **대상**(**확률변수**)에 대한 **확률** 또는 **확률분포**를 **계산**할 수 있는 모델
- **확률분포**를 이용한 **지식표현**을 하고 **확률적 추론**

- **베이지안 망**
- **마르코프 랜덤 필드(마르코프 망)**
- **조건부 랜덤 필드**
- **로그-선형 모델**

확률 그래프 모델

- ❖ **확률 그래프 모델(probabilistic graphical model)**
  - ▪ **예. 절도 경보 문제**
    - **절도**가 발생하거나 **지진**이 발생하면 **경보** 발생
    - 경보가 울리면 **이웃**이 **전화**

    - **불확실한 요소**가 있어 **확률**로 표현
      - **확률변수**(random variable)
        - » 경보 작동(**A**; alarm)
        - » 절도 발생(**B**; burglary)
        - » 지진 발생(**E**; earthquake)
        - » 이웃 전화(**N**; neighbor call)

확률 그래프 모델

- **절도 경보 문제의 확률분포에 의한 지식표현**
  - 지진 발생(E), 절도 발생(B), 경보 생성(A), 이웃전화 (N)
  - **결합확률 분포**로 표현

E	B	A	N	확률
F	F	F	F	0.56133
F	F	F	T	0.06237
F	F	T	F	0.00126
F	F	T	T	0.00504
F	T	F	F	0.0243
F	T	F	T	0.0027
F	T	T	F	0.0486
F	T	T	T	0.1944
T	F	F	F	0.0189
T	F	F	T	0.0021
T	F	T	F	0.0098
T	F	T	T	0.0392
T	T	F	F	0.00027
T	T	F	T	0.00003
T	T	T	F	0.00594
T	T	T	T	0.02376

경보가 울릴 때 이웃이 전화할 확률은?

$$P(N = T | A = T) = ?$$

이웃이 전화했을 때 도둑이 들었을 확률은?

$$P(B = T | N = T) = ?$$

확률 그래프 모델

❖ 조건부 독립과 확률분포의 인수분해(factorization)

- 사건의 **독립**(independence)
  - $P(E,B) = P(E)P(B)$

- **조건부 독립**(conditional independence) 성질 이용
  - $P(N,A|E) = P(N|E)P(A|E)$

- **확률분포**의 **인수분해**
  - $P(A,B) = P(A|B)P(B)$
  - $P(A_1,A_2,A_3,A_4) = P(A_1|A_2,A_3,A_4)P(A_2|A_3,A_4)P(A_3|A_4)P(A_4)$

확률 그래프 모델

❖ 조건부 독립을 이용한 확률분포의 인수 분해

- $P(N, A, E, B) = P(N|A, E, B)P(A|E, B)P(E|B)P(B)$

E	B	A	N	확률
F	F	F	F	0.56133
F	F	F	T	0.06237
F	F	T	F	0.00126
F	F	T	T	0.00504
F	T	F	F	0.0243
F	T	F	T	0.0027
F	T	T	F	0.0486
F	T	T	T	0.1944
T	F	F	F	0.0189
T	F	F	T	0.0021
T	F	T	F	0.0098
[illegible]	[illegible]	T	T	0.0392
[illegible]	[illegible]	F	F	0.00027
[illegible]	[illegible]	F	T	0.00003
[illegible]	[illegible]	T	F	0.00594
[illegible]	[illegible]	T	T	0.02376

만족하는 조건부 독립 성질

$P(N|A, E, B) = P(N|A)$

$P(E|B) = P(E)$

$$= P(N|A)P(A|E, B)P(E)P(B)$$

A	N: F	N: T
F	0.9	0.1
T	0.2	0.8

E	B	A: F	A: T
F	F	0.99	0.01
F	T	0.1	0.9
T	F	0.3	0.7
T	T	0.01	0.99

E: F	E: T
0.9	0.1

B: F	B: T
0.7	0.3

$$P(N = T, A = T, E = F, B = T)$$

$$= P(N = T|A = T)P(A = T|E = F, B = T)P(E = F)P(B = T)$$

$$= 0.8 \times 0.9 \times 0.9 \times 0.3$$

$$= 0.1944$$

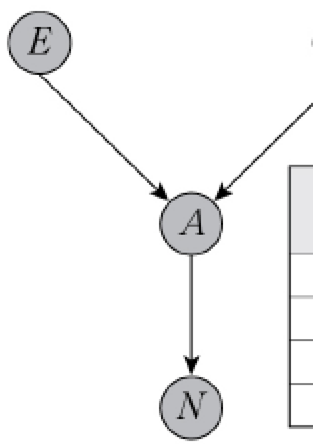
2. 베이지안 망

❖ **베이지안 망**(**Bayesian network**)

- **확률변수** 간의 **조건부 독립**을 표현한 **방향성 그래프**(directed graph)와, **조건부 확률분포**들로 확률분포를 표현한 것
  - **노드** : 확률 변수
  - **간선** : 의존관계
- $P(N,A,E,B) = P(N|A,E,B)P(A|E,B)P(E|B)P(B) = P(N|A)P(A|E,B)P(E)P(B)$

E	
F	T
0.9	0.1

B	
F	T
0.7	0.3



E	B	A	
		F	T
F	F	0.99	0.01
F	T	0.1	0.9
T	F	0.3	0.7
T	T	0.01	0.99

A	N	
	F	T
F	0.9	0.1
T	0.2	0.8

베이지안 망

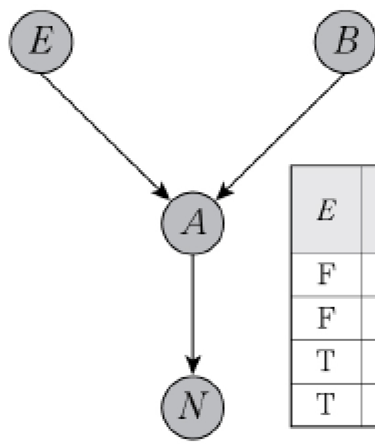
❖ **베이지안 망** - cont.

- $P(N,A,E,B) = P(N|A,E,B)P(A|E,B)P(E|B)P(B)$

$$= P(N|A)P(A|E,B)P(E)P(B)$$

E	
F	T
0.9	0.1

B	
F	T
0.7	0.3



E	B	A	
		F	T
F	F	0.99	0.01
F	T	0.1	0.9
T	F	0.3	0.7
T	T	0.01	0.99

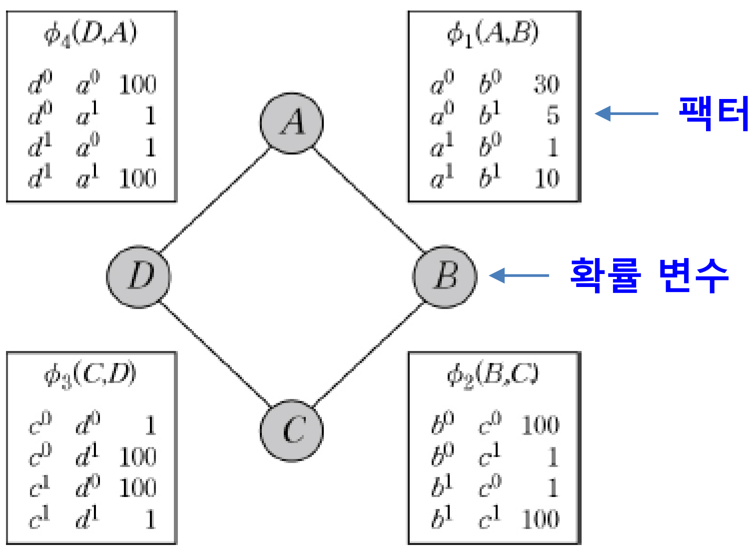
A	N	
	F	T
F	0.9	0.1
T	0.2	0.8

$$P(N=\mathrm{F}, A=\mathrm{F}, E=\mathrm{F}, B=\mathrm{F})$$
$$= P(N=\mathrm{F}|A=\mathrm{F})P(A=\mathrm{F}|E=\mathrm{F}, B=\mathrm{F})P(E=\mathrm{F})P(B=\mathrm{F})$$
$$= 0.9 \cdot 0.99 \cdot 0.9 \cdot 0.7 = 0.56133$$

3. 마르코프 랜덤 필드

❖ **마르코프 랜덤 필드(Markov random field)**

- **마르코프 망**(Markov network)라고도 함
- 확률분포를 **무방향 그래프**(undirected graph)를 사용하여 표현
- 확률변수들의 값의 조합에 대한 값을 부여한 **팩터**(factor, potential function)
  - 각 조합에 대한 **호응 정도**(affinity, compatibility) 정의
    - 호응 정도는 0이상의 실수 값

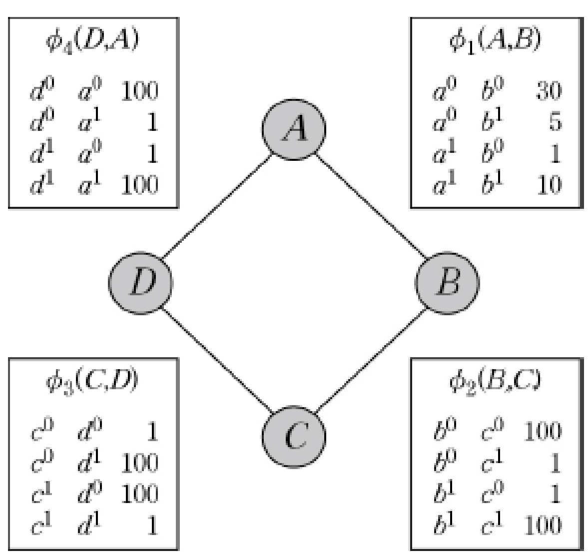


마르코프 랜덤 필드

❖ **마르코프 랜덤 필드(Markov network)** – cont.

- 확률변수들의 **값의 조합**에 대한 값을 부여한 **팩터(factor)들의 곱**에 **비례**하는 **확률값** 표현
- **분할함수**(partition function) **값**
  - 팩터 곱들의 전체 합
- **확률 = (팩터의 곱)/(분할함수의 값)**

$$P(a^0, b^0, c^0, d^0) = 300{,}000/7{,}201{,}840$$
$$= 0.04$$



A	B	C	D	팩터의 곱	확률
a^0	b^0	c^0	d^0	300,000	0.04
a^0	b^0	c^0	d^1	300,000	0.04
a^0	b^0	c^1	d^0	300,000	0.04
a^0	b^0	c^1	d^1	30	$4.1 \cdot 10^{-6}$
a^0	b^1	c^0	d^0	500	$6.9 \cdot 10^{-5}$
a^0	b^1	c^0	d^1	500	$6.9 \cdot 10^{-5}$
a^0	b^1	c^1	d^0	5,000,000	0.69
a^0	b^1	c^1	d^1	500	$6.9 \cdot 10^{-5}$
a^1	b^0	c^0	d^0	100	$1.4 \cdot 10^{-5}$
a^1	b^0	c^0	d^1	1,000,000	0.14
a^1	b^0	c^1	d^0	100	$1.4 \cdot 10^{-5}$
a^1	b^0	c^1	d^1	100	$1.4 \cdot 10^{-5}$
a^1	b^1	c^0	d^0	10	$1.4 \cdot 10^{-6}$
a^1	b^1	c^0	d^1	100,000	0.014
a^1	b^1	c^1	d^0	100,000	0.014
a^1	b^1	c^1	d^1	100,000	0.014

팩터곱의 합(분할함수의 값): 7, 201, 840

확률 그래프 모델

- ❖ **연속인 확률변수가 포함된 확률 분포**
  - 표(table)를 사용한 표현 곤란
  - **함수식**을 이용한 표현
    - **베이지안 망** : 조건부 확률 값 출력 **함수** 사용
    - **마르코프 랜덤 필드** : **지수함수**와 같은 **함수**로 팩터 정의

$\phi(X_i, X_j, X_k) = \exp(f(X_i, X_j, X_j))$: **팩터**(factor)

X_i, X_j, X_k : 확률 변수

$f(X_i, X_j, X_k)$: 특정 **특징**의 **유무**나 **정도** 등을 계산하는 **함수식**

예. $f(X_1, X_2, X_3) = 2X_1 - 3X_2X_3$

$f(X_1, X_2) = X_1, X_2$가 공통으로 겹치는 정도

$f(X_1) = X_1$을 알고리즘 A로 처리한 결과값

$$f(X_1, X_2) = \begin{cases} 1.5 & \text{if } X_1 = X_2 \\ 0.1 & \text{otherwise} \end{cases}$$

$$f(X_1, X_2) = \begin{cases} 0.2 & \text{if } X_1 = X_2 \\ 1.3 & \text{otherwise} \end{cases}$$

4. 조건부 랜덤 필드

❖ **조건부 랜덤 필드**(conditional random field, **CRF**)

- **조건부 확률분포**를 **표현**하는 마르코프 랜덤 필드

- $\boldsymbol{X}$: **관측되는 대상**이나 **입력**을 나타내는 확률변수들의 집합
- $\boldsymbol{Y}$: **추정**하거나 **예측**하는 **대상**을 나타내는 확률변수들의 집합
- 관측값 $\boldsymbol{X}$가 주어질 때 $\boldsymbol{Y}$의 확률계산
- $\{\phi_1(D_1), \phi_2(D_2), \cdots, \phi_n(D_n)\}$: **Y**의 확률변수를 **하나라도 포함한 팩터**의 집합

- **조건부 확률 정의**

$$P(\boldsymbol{Y}|\boldsymbol{X}) = \frac{1}{Z(\boldsymbol{X})}\tilde{P}(\boldsymbol{X}, \boldsymbol{Y})$$

$$\tilde{P}(\boldsymbol{X}, \boldsymbol{Y}) = \prod_{i=1}^{m}\phi_i(D_i)$$

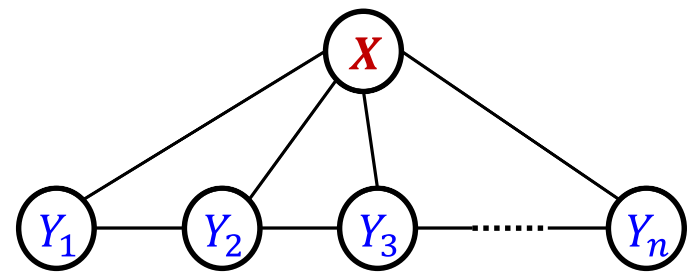
$$Z(\boldsymbol{X}) = \sum_{\boldsymbol{Y}}\tilde{P}(\boldsymbol{Y}, \boldsymbol{X})$$ 분할 함수(partition function)

조건부 랜덤 필드

❖ 조건부 랜덤 필드의 예

- 자연어 품사 태깅

$\boldsymbol{X} = \{X_1, X_2, \dots, X_n\}$



	X_1	X_2	X_3	X_4	X_5	X_6
입력	**The**	**boy**	**knocks**	**at**	**the**	**door.**
출력	**D**	**N**	**V**	**P**	**D**	**N**
	Y_1	Y_2	Y_3	Y_4	Y_5	Y_6

$$P(Y|X) = \frac{1}{Z}\exp\left(\sum_j \sum_{i=1}^{n-1} w_i f_j(Y_{i+1}, Y_i, \boldsymbol{X}, i) + \sum_j \sum_{i=1}^{n-1} v_i g_j(Y_i, \boldsymbol{X}, i)\right)$$

$$f_j(Y_{i+1}, Y_i, \boldsymbol{X}, i) = \begin{cases} 1 & \text{if } Y_{i+1} = P,\ Y_i = V, \text{and } X_i = \text{knocks} \\ 0 & \text{otherwise} \end{cases}$$

$$g_j(Y_i, \boldsymbol{X}, i) = \begin{cases} 1 & \text{if } Y_i = V, \text{and } X_i = \text{knocks} \\ 0 & \text{otherwise} \end{cases}$$

5. 로그-선형 모델

- **로그-선형 모델**(log-linear model)
  - **팩터**가 **지수 함수**로 **표현**되는 **마르코프 랜덤 필드 모델**
    - $\boldsymbol{X} = \{X_1, X_2, \cdots, X_m\}$: 확률변수의 집합
    - **팩터**의 형태

$$\phi_i = \exp(-w_i f_i(D_i))$$

계수(parameter)

확률변수 집합 D_i에 정의된 함수: 특징 추출

- **확률 분포**의 표현

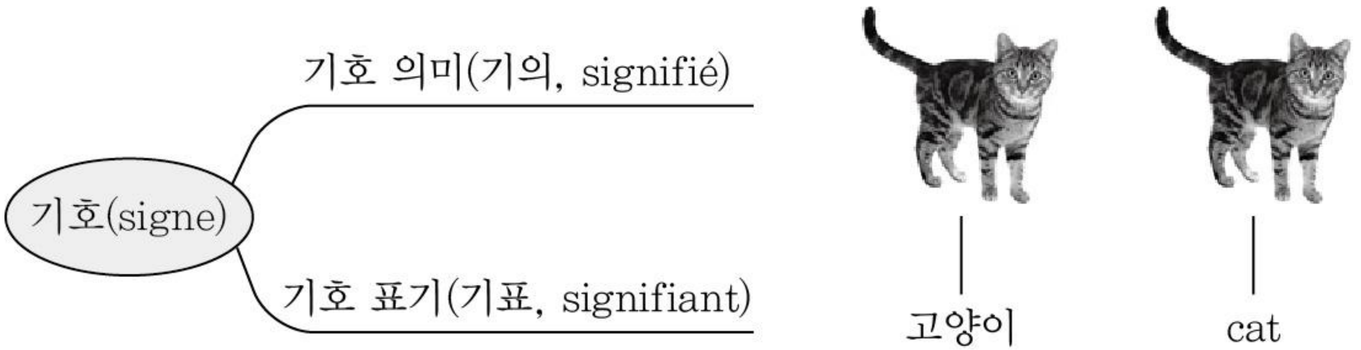
$$P(X_1, X_2, \cdots, X_n) = \frac{1}{Z} exp\left[-\sum_{i=1}^{K} w_i f_i(D_i)\right]$$

분할 함수(partition function)

6. 심볼 그라운딩 문제와 프레임 문제

❖ 심볼 그라운딩 문제

- '**고양이는 귀엽다**'라는 지식 표현
  - '고양이'와 '귀엽다'는 대상 또는 개념을 가리키는 **기호**(symbol) 사용



- 기호의 **표기**와 **의미**가 **자의적인 관계**
- 기호를 이해하는 **문화 체계** 속에서 필연화

- **심볼 그라운딩**(symbol grounding)
  - **기호 표기**를 실제 세계의 **의미**와 **연결**시키는 것

심볼 그라운딩 문제

❖ **심볼 그라운딩 문제 – cont.**

- **기호 표기**로 표현되어 있는 **지식**에 대해서, 컴퓨터는 **심볼 그라운딩**을 할 수 있는 **능력**이 **없음**
- 실제 세계와 컴퓨터의 기호 표기 사이의 심볼 그라운딩을 **인간**이 **대신 수행**
- 기호 표기를 실제 세계의 의미와 직접 연결시킬 수 없다는 것
- 딥러닝 기술 발전은 심볼 그라운딩 문제에 해결에 기여 예상

프레임 문제

❖ 프레임(frame) 문제

- **사고범위 문제(思考範圍問題)**
- 어떤 작업을 수행할 때 **관련 있는 지식**만 꺼내서 사용한다는 것은 지극히 자연스럽고 당연하지만, 인공지능에서는 이러한 일이 쉽지 않다는 것

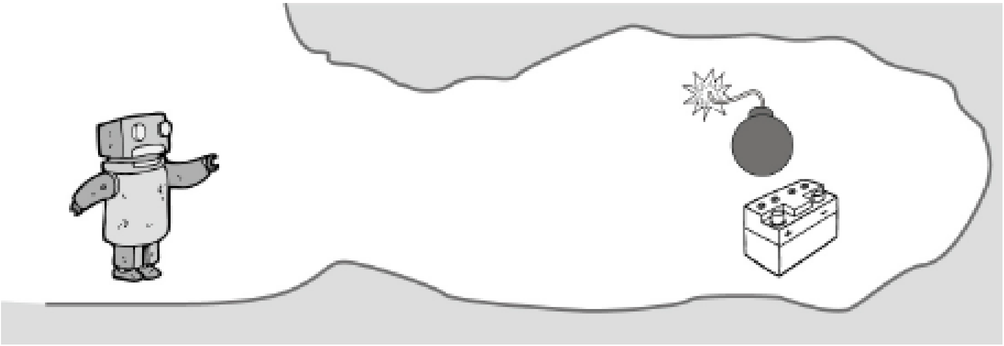


그림 3.28 프레임 문제

7. CYC 프로젝트

❖ 상식의 필요성

- 추론 등을 위해 **상식**(commonsense)의 활용 중요
  - 상식의 예
    - 물체를 공중에서 놓으면 아래로 떨어진다.
    - 사람은 태어나기 전에는 존재하지 않는다.
    - 물고기는 물에서 살며 물 밖으로 나오면 죽는다.
    - 빵은 빵가게에서 산다.
    - 물컵이 넘어지면 물이 나온다.
  - 상식의 적용 예 - 기계번역

He saw a girl in the garden with a telescope.

그는 망원경으로 정원에 있는 소녀를 보았다.
그는 정원에서 망원경으로 소녀를 보았다.
그는 소녀가 정원에서 망원경을 들고 있는 것을 보았다.
정원에서 그는 망원경을 들고 있는 소녀를 보았다.

CYC 프로젝트

❖ CYC 프로젝트

- **상식적인 추론**을 하는데 필요한 방대한 **지식**을 **추출**하여 **표현**하는 프로젝트
- **일차 술어 논리**를 사용 지식 표현
- 1984년 르냇(Douglas Lenat) 시작

- 예. CYC의 예

(#$isa #$DonaldTrump #$UnitedStatesPresident) ; 도날드트럼프는 미국 대통령이다.
(#$capitalCity #$France #$Paris) ; 프랑스 수도는 파리이다.
(#$implies (#$and (#$isa ?OBJ ?SUBSET) (#$genls ?SUBSET ?SUPERSET))
(#$isa ?OBJ ?SUPERSET)) ; OBJ가 SUBSET의 사례이고, SUBSET이
SUPERSET에 속하면, OBJ는 SUPERSET의 사례이다. **(규칙표현의 예)**
(#$relationAllExists #$biologicalMother #$ChordataPhylum #$FemaleAnimal)
; 모든 ChordataPhylum에 속하는 개체에게는 어머니(biologicalMother)인
; 여성(FemaleAnimal)이 있다. **(존재한정사가 있는 문장)**

[실습] 확률 그래프 모델

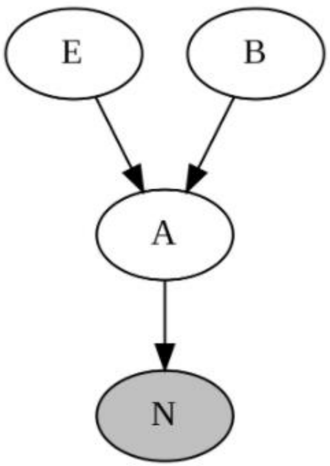
❖ pgmpy 패키지 사용

- 다양한 확률 그래프 모델 지원

```
!pip install pgmpy
```

❖ 베이지안 망 정의 및 추론

```
import networkx as nx
from networkx.drawing.nx_pydot import to_pydot
from IPython.core.display import Image

g = nx.DiGraph()
g.add_edge("E", "A")
g.add_edge("B", "A")
g.add_edge("A", "N")
d = to_pydot(g)
d.get_node("N")[0].set_fillcolor("gray")
d.get_node("N")[0].set_style("filled")
d.set_dpi(300)
d.set_margin(0.2)
Image(d.create_png(), width=200)
```



```python
from pgmpy.factors.discrete import TabularCPD
import numpy as np

# 지진(Earthquake) 발생 확률 분포
P_E = TabularCPD('E', 2, [[0.9], [0.1]], state_names={'E': ['F', 'T']})
print('P(E)')
print(P_E)

# 절도(Burglary) 발생 확률분포
P_B = TabularCPD('B', 2, [[0.7], [0.3]], state_names={'B': ['F', 'T']})
print('P(B)')
print(P_B)

# 경보(Alarm) 발생 확률 분포
P_A_I_EB = TabularCPD('A', 2, [[0.99, 0.1, 0.3, 0.01], [0.01, 0.9, 0.7, 0.99]],
                      evidence=['E', 'B'], evidence_card=[2, 2],
                      state_names={'A': ['F','T'],'E': ['F','T'],'B':['F','T']})
print('P(A|EB)')
print( P_A_I_EB)

# 이웃(Neighbor) 전화 확률 분포
P_N_I_A = TabularCPD('N', 2,
    np.array([[0.9, 0.2], [0.1,0.8]]),
    evidence=['A'], evidence_card=[2],
    state_names={'N': ['F', 'T'],'A': ['F', 'T']})
print('P(N|A)')
print(P_N_I_A)
```

```
P(E)
+------+-----+
| E(F) | 0.9 |
+------+-----+
| E(T) | 0.1 |
+------+-----+

P(B)
+------+-----+
| B(F) | 0.7 |
+------+-----+
| B(T) | 0.3 |
+------+-----+

P(A|EB)
+------+------+------+------+------+
| E    | E(F) | E(F) | E(T) | E(T) |
+------+------+------+------+------+
| B    | B(F) | B(T) | B(F) | B(T) |
+------+------+------+------+------+
| A(F) | 0.99 | 0.1  | 0.3  | 0.01 |
+------+------+------+------+------+
| A(T) | 0.01 | 0.9  | 0.7  | 0.99 |
+------+------+------+------+------+

P(N|A)
+------+------+------+
| A    | A(F) | A(T) |
+------+------+------+
| N(F) | 0.9  | 0.2  |
+------+------+------+
| N(T) | 0.1  | 0.8  |
+------+------+------+
```

```python
from pgmpy.models import BayesianModel

# 베이지안 망 구조 정의
model = BayesianModel([('E','A'), ('B','A'),('A','N')])
model.add_cpds(P_E, P_B, P_A_l_EB, P_N_l_A)  # 확률분포 등록
model.check_model()
```

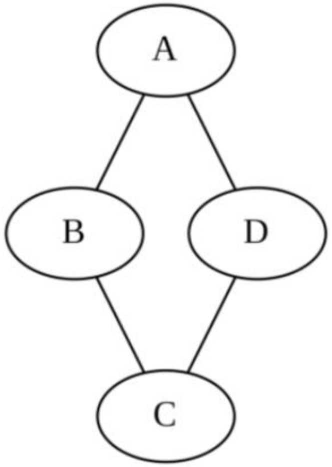
```
True
```

```python
from pgmpy.inference import VariableElimination

# 베이지안 망의 추론
infer = VariableElimination(model)
A_dist = infer.query(['A'])
print('P(A)')
print(A_dist)

N_l_EF_BT = infer.query(['N'], evidence={'E':'F', 'B':'T'})
print('P(N | E=F,B=T)')
print(N_l_EF_BT)

N_l_AF_BT = infer.query(['N'], evidence={'A':'F', 'B':'T'})
print('P(N | A=F,B=T)')
print(N_l_AF_BT)
```

```
+------+----------+
| A    |   phi(A) |
+======+==========+
| A(F) |   0.6720 |
+------+----------+
| A(T) |   0.3280 |
+------+----------+

P(N | E=F,B=T)
+------+----------+
| N    |   phi(N) |
+======+==========+
| N(F) |   0.2700 |
+------+----------+
| N(T) |   0.7300 |
+------+----------+

P(N | A=F,B=T)
+------+----------+
| N    |   phi(N) |
+======+==========+
| N(F) |   0.9000 |
+------+----------+
| N(T) |   0.1000 |
+------+----------+
```

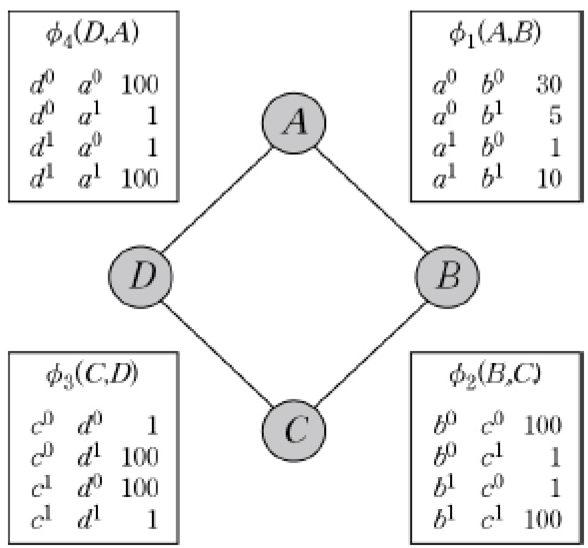
❖ 마르코프 랜덤 필드의 예

```python
import networkx as nx
from IPython.core.display import Image
from networkx.drawing.nx_pydot import to_pydot

g1 = nx.Graph()
g1.add_edge("A", "B")
g1.add_edge("D", "A")
g1.add_edge("B", "C")
g1.add_edge("C", "D")

d1 = to_pydot(g1)
d1.set_dpi(300)
d1.set_margin(0.5)
Image(d1.create_png(), width=300)
```



```python
from pgmpy.models import MarkovModel
from pgmpy.factors.discrete import DiscreteFactor

# 마르코프 랜덤 필드(마르코프 모델) 모델 정의
model = MarkovModel([('A', 'B'), ('B', 'C'), ('C','D'), ('D','A')])
factor1 = DiscreteFactor(['A', 'B'], [2, 2], [30, 5, 1, 10], state_names={'A': [0,1],'B': [0,1]})
factor2 = DiscreteFactor(['B', 'C'], [2, 2], [100, 1, 1, 100], state_names={'B': [0,1],'C': [0,1]})
factor3 = DiscreteFactor(['C', 'D'], [2, 2], [1, 100, 100, 1], state_names={'C': [0,1],'D': [0,1]})
factor4 = DiscreteFactor(['D', 'A'], [2, 2], [100, 1, 1, 100], state_names={'D': [0,1],'A': [0,1]})
model.add_factors(factor1, factor2, factor3, factor4)
print('모델의 타당성: ', model.check_model())
```

```
모델의 타당성:  True
```



```python
import numpy as np
pf_value = model.get_partition_function()
print('₩n분할 함수의 값: ', pf_value)

infer = VariableElimination(model)  # 추론 객체 생성

phi_ABCD = infer.query(['A', 'B','C', 'D'])   # 전체 분포 phi(A,B,C,D)
print('phi(A,B,C,D)')
print(phi_ABCD)
P_ABCD = phi_ABCD.values/pf_value   # 확률 = (팩터의 곱)/(분할 함수의 값)
PABCD = np.reshape(P_ABCD, -1)
for val in PABCD: # 확률의 출력
  print(val, '₩n')

A_dist = infer.query(['A'])  # A의 분포 phi(A)
print('phi(A)')
print(A_dist)
P_A = A_dist.values/np.sum(A_dist.values)
for val in P_A:
  print(val, '₩n')

AlB0C1_dist = infer.query(['A'], evidence={'B':0, 'C':1}) # phi(A|B=0,C=1)
print('phi(A|B=0,C=1)')
print(AlB0C1_dist)
P_AlB0C1 = AlB0C1_dist.values/np.sum(AlB0C1_dist.values)
for val in P_AlB0C1:
  print(val, '₩n')
```

```
분할 함수의 값:  7201840.0
phi(A,B,C,D)
+------+------+------+------+----------------+
| C    | D    | A    | B    |   phi(C,D,A,B) |
+======+======+======+======+================+
| C(0) | D(0) | A(0) | B(0) |    300000.0000 |   0.0416560212390167
+------+------+------+------+----------------+
| C(0) | D(0) | A(0) | B(1) |       500.0000 |   6.942670206502783e-05
+------+------+------+------+----------------+
| C(0) | D(0) | A(1) | B(0) |       100.0000 |   1.3885340413005566e-05
+------+------+------+------+----------------+
| C(0) | D(0) | A(1) | B(1) |        10.0000 |   1.3885340413005566e-06
+------+------+------+------+----------------+
| C(0) | D(1) | A(0) | B(0) |    300000.0000 |   0.0416560212390167
+------+------+------+------+----------------+
| C(0) | D(1) | A(0) | B(1) |       500.0000 |   6.942670206502783e-05
+------+------+------+------+----------------+
| C(0) | D(1) | A(1) | B(0) |   1000000.0000 |   0.13885340413005565
+------+------+------+------+----------------+
| C(0) | D(1) | A(1) | B(1) |    100000.0000 |   0.013885340413005565
+------+------+------+------+----------------+
| C(1) | D(0) | A(0) | B(0) |    300000.0000 |   0.0416560212390167
+------+------+------+------+----------------+
| C(1) | D(0) | A(0) | B(1) |   5000000.0000 |   0.6942670206502782
+------+------+------+------+----------------+
| C(1) | D(0) | A(1) | B(0) |       100.0000 |   1.3885340413005566e-05
+------+------+------+------+----------------+
| C(1) | D(0) | A(1) | B(1) |    100000.0000 |   0.013885340413005565
+------+------+------+------+----------------+
| C(1) | D(1) | A(0) | B(0) |        30.0000 |   4.16560212390167e-06
+------+------+------+------+----------------+
| C(1) | D(1) | A(0) | B(1) |       500.0000 |   6.942670206502783e-05
+------+------+------+------+----------------+
| C(1) | D(1) | A(1) | B(0) |       100.0000 |   1.3885340413005566e-05
+------+------+------+------+----------------+
| C(1) | D(1) | A(1) | B(1) |    100000.0000 |   0.013885340413005565
+------+------+------+------+----------------+
```

```
phi(A)
+------+--------------+
| A    |       phi(A) |
+======+==============+
| A(0) | 5901530.0000 |   0.8194475300756473
+------+--------------+
| A(1) | 1300310.0000 |   0.18055246992435267
+------+--------------+
```

```
phi(A|B=0,C=1)
+------+-------------+
| A    |      phi(A) |
+======+=============+
| A(0) | 300030.0000 |   0.9993338440528928
+------+-------------+
| A(1) |    200.0000 |   0.0006661559471072178
+------+-------------+
```

Quiz

❖ **확률 분포에 대한 설명으로 적합하지 않는 것을 선택하시오.**

① 관심대상을 표현하는 확률변수들에 대한 결합확률 분포를 가지고 있으면 다양한 상황에 대한 확률적 추론을 할 수 있다.

② 결합확률 분포를 가지고 있으면 특정한 조건부 확률이나 일부 확률변수에 대한 결합확률 분포를 계산할 수 있다.

③ 확률분포를 사용하여 불확실한 사건이나 지식을 표현할 수 있다.

④ 확률변수들은 서로 독립적이기 때문에 어떤 확률변수로 다른 확률변수에 대한 정보를 유추하는 것은 불가능하다.

❖ **확률에 대한 설명으로 적합하지 않는 것을 선택하시오.**

① 확률변수 A가 갖는 값이 확률변수 B가 갖는 값과 전혀 연관이 없을 때 이들 확률변수는 서로 독립이다고 한다.

② 모든 확률변수가 서로 독립이면, 결합확률 분포는 확률분포들의 곱으로 표현할 수 없다.

③ 특정 확률변수 C의 값이 주어지면 확률변수 A와 B가 서로 독립일 때, C가 주어질 때 A와 B는 조건부 독립이라고 한다.

④ 조건부 독립의 성질을 갖는 확률변수들이 있으면 결합확률 분포를 확률분포들의 곱으로 표현할 수 있다.

Quiz

❖ **확률 그래프 모델에 대한 설명으로 적합하지 않는 것을 선택하시오**.

① 확률 그래프 모델은 그래프 구조를 이용하여 확률변수간의 연관관계를 표현하여 확률분포에 표현 및 추론을 용이하게 한다.

② 마르코프 랜덤 필드에서 팩터는 지수함수로 정의해야 한다.

③ 마르코프 랜덤 필드에서는 확률변수들의 값의 조합별로 서로 부합되는 정도를 나타내는 호응정도값을 부여하는 팩터를 결정하고, 이 팩터들의 곱을 사용하여 확률값을 결정한다.

④ 베이지안 망에서는 방향 그래프를 사용하여 조건부 독립의 관계를 표현한다.

❖ **다음 지식표현에 관련한 설명으로 적합하지 않는 것을 선택하시오.**

① 일상 상황에 대한 문제 해결을 위해서는 상식에 대한 지식이 필수적이며, 상식을 일차 술어 논리 형태의 지식을 구축하여, 이러한 문제 해결은 비교적 쉽게 해결할 수 있다

② 많은 지식을 포함하고 있는 시스템에 현재 해결하려는 문제에 관련된 지식을 선택하는 과정에서 시간이 너무 오래 걸릴 수 있는 문제를 프레임 문제라고 한다.

③ 기호 표기를 실제 세계의 의미와 연결시키는 것을 심볼 그라운딩이라 한다.

④ 기호로 지식을 표현하는 컴퓨터는 심볼 그라운딩을 할 수 있는 능력이 아직은 없다.