In [1]:
```
import numpy as np
from pomegranate import *
```

1.

Write a code for building the Bayesian network diagram above using Pomegranate (or using whatever you want).

Sol)

In [2]:
```
# 노드 구성

Smokes = Node(DiscreteDistribution({
    'smokesTrue' : 0.2,
    'smokesFalse' : 0.8,
}),name='Smokes')

LungDisease = Node(ConditionalProbabilityTable([
    ['smokesTrue','lungTrue',0.1],
    ['smokesTrue','lungFalse',0.9],
    ['smokesFalse','lungTrue',0.001],
    ['smokesFalse','lungFalse',0.999],
],[Smokes.distribution()],name='LungDisease')

ShortnessOfBreath = Node(ConditionalProbabilityTable([
    ['lungTrue','shortTrue',0.2],
    ['lungTrue','shortFalse',0.8],
    ['lungFalse','shortTrue',0.01],
    ['lungFalse','shortFalse',0.99],
],[LungDisease.distribution()],name='ShortnessOfBreath')

ChestPain = Node(ConditionalProbabilityTable([
    ['lungTrue','chestTrue',0.2],
    ['lungTrue','chestFalse',0.8],
    ['lungFalse','chestTrue',0.01],
    ['lungFalse','chestFalse',0.99],
],[LungDisease.distribution()],name='ChestPain')

Cold = Node(DiscreteDistribution({
    'coldTrue' : 0.02,
    'coldFalse' : 0.98,
}),name='Cold')

Fever = Node(ConditionalProbabilityTable([
    ['coldTrue','feverTrue',0.3],
    ['coldTrue','feverFalse',0.7],
    ['coldFalse','feverTrue',0.01],
    ['coldFalse','feverFalse',0.99],
],[Cold.distribution()],name='fever')

Cough = Node(ConditionalProbabilityTable([
    ['lungTrue','coldTrue','coughTrue',0.75],
    ['lungTrue','coldTrue','coughFalse',0.25],
    ['lungTrue','coldFalse','coughTrue',0.5],
    ['lungTrue','coldFalse','coughFalse',0.5],
    ['lungFalse','coldTrue','coughTrue',0.5],
    ['lungFalse','coldTrue','coughFalse',0.5],
    ['lungFalse','coldFalse','coughTrue',0.01],
    ['lungFalse','coldFalse','coughFalse',0.99],
],[LungDisease.distribution(), Cold.distribution()],name='Cough')


# 모델 생성

model = BayesianNetwork()
model.add_states(Smokes,LungDisease,ShortnessOfBreath,ChestPain,Cough,Fever,Cold)

# edge 추가

model.add_edge(Smokes, LungDisease)
model.add_edge(LungDisease, ShortnessOfBreath)
model.add_edge(LungDisease, ChestPain)
model.add_edge(LungDisease, Cough)
model.add_edge(Cold, Cough)
model.add_edge(Cold, Fever)

# 완료

model.bake()
```

2.

Write a code for calculating the joint probability

$P(Smokes = T, LungDisease = T, Shortness\ of\ Breath = T, ChestPain = F, Cough = T, Fever = F, Cold = F)$. (Attach the code and result)

Sol)

In [15]:
```
prob = model.probability([['smokesTrue','lungTrue','shortTrue','chestFalse','coughTrue','feverFalse','coldFalse']])

print(f'P(Smokes = T, LungDisease = T, Shortness of Breath = T, ChestP ain = F, Cough = T, Fever = F, Cold = F) = {prob:.8f}')
```

```
P(Smokes = T, LungDisease = T, Shortness of Breath = T, ChestP ain = F, Cough = T, Fever = F, Cold = F) = 0.00155232
```

3.

Write the equation and explanation for calculating the probability of the Problem 2. (You can use the calculator)

Sol)

$P(Smokes = T, LungDisease = T, Shortness\ of\ Breath = T, ChestPain = F, Cough = T, Fever = F, Cold = F)$

$= P(ShortnessofBreath = T|LungDisease = T) \times P(ChestPain = False|LungDisease = True) \times$

$P(Cough = T|LungDisease = T, Cold = F) \times P(Fever = F|Cold = F) \times P(LungDisease = T|Smokes = T) \times$

$P(Smokes = T) \times P(Cold = False)$

= 0.2 × 0.8 × 0.5 × 0.99 × 0.1 × 0.2 × 0.98

= 0.00155232

4.

Write a code for calculating the probabilities of all nodes given ChestPain = T.

For example, calculate P(Smokes = T|ChestPain = T), P(Smokes = F|ChestPain = T), P(LungDisease = T|ChestPain = T),... (Attach the code and result)

Sol)

In [14]:
```
# pomegranate의 조건부확률 코드를 사용하자.

predictions = model.predict_proba({
    "ChestPain": "chestTrue",
})


# 출력하는 코드
for node, prediction in zip(model.states, predictions):
    if isinstance(prediction, str):
        print(f"{node.name}: {prediction}")
    else:
        print(f"{node.name}")
        for value, probability in prediction.parameters[0].items():
            if value[:2] == 'no':
                valuetemp = value[0:-5]
                print(f"  P({valuetemp} = T|ChestPain = T): {probability:.4f}")
            if value[:2] == 'sm':
                valuetemp = value[0:-4]
                print(f"  P({valuetemp} = T|ChestPain = T): {probability:.4f}")
```

```
Smokes
  P(smokes = T|ChestPain = T): 0.4157
  P(smokes = F|ChestPain = T): 0.5843
LungDisease
  P(lung = F|ChestPain = T): 0.7018
  P(lung = T|ChestPain = T): 0.2982
ShortnessOfBreath
  P(short = F|ChestPain = T): 0.9333
  P(short = T|ChestPain = T): 0.0667
ChestPain: chestTrue
Cough
  P(cough = T|ChestPain = T): 0.1645
  P(cough = F|ChestPain = T): 0.8355
Cold
  P(fever = F|ChestPain = T): 0.9842
  P(fever = T|ChestPain = T): 0.0158
Cold
  P(cold = T|ChestPain = T): 0.0200
  P(cold = F|ChestPain = T): 0.9800
```

5.

Write the equation and explanation for calculating the probability $P(Cough = T|ChestPain = T)$ of the Problem 4. (You can use the calculator)

Sol)

$P(Cough = T|ChestPain = T) \propto P(Cough = T, ChestPain = T)$ ($\propto$ = normalize factor)

$= \sum_{Smokes,Cold,LungDisease} \alpha P(Cough = T, ChestPain = T, \mathbf{Smokes, Cold, LungDisease})$ (Bold letters are variable.)

$= \sum_{Smokes,Cold,LungDisease} \alpha P(\mathbf{Smokes})P(\mathbf{Cold})P(\mathbf{LungDisease|Smokes})P(ChestPain = T|\mathbf{LungDisease})P(Cough = T|\mathbf{Cold, LungDisease})$

α(0.2 × 0.02 × 0.1 × 0.2 × 0.75 (TTT)

+0.2 × 0.02 × 0.9 × 0.01 × 0.5 (TTF)

+0.2 × 0.98 × 0.1 × 0.2 × 0.5 (TFT)

+0.2 × 0.98 × 0.9 × 0.01 × 0.01 (TFF)

+0.8 × 0.02 × 0.001 × 0.2 × 0.75 (FTT)

+0.8 × 0.02 × 0.999 × 0.01 × 0.5 (FTF)

+0.8 × 0.98 × 0.001 × 0.2 × 0.5 (FFT)

+0.8 × 0.98 × 0.999 × 0.01 × 0.01) (FFF)

$= \alpha \times 0.0022946816$

In [5]:
```
# 확률 계산

a = [0.2,0.2,0.2,0.2,0.8,0.8,0.8,0.8]
b = [0.02,0.02,0.98,0.98,0.02,0.02,0.98,0.98]
c = [0.1,0.9,0.1,0.9,0.001,0.999,0.001,0.999]
d = [0.2,0.8,0.2,0.8,0.2,0.8,0.2,0.8]
e = [0.75,0.5,0.5,0.01,0.75,0.5,0.5,0.01]
sol1 = 0
for i in range(8):
    sol1 += a[i]*b[i]*c[i]*d[i]*e[i]

print(f'{sol1:.10f}')
```

```
0.0022946816
```

Find $P(cough = F|ChestPain = T)$ in the same way.

$P(Cough = F|ChestPain = T) \propto P(Cough = F, ChestPain = T)$ ($\propto$ = normalize factor)

$= \sum_{Smokes,Cold,LungDisease} \alpha P(Cough = F, ChestPain = T, \mathbf{Smokes, Cold, LungDisease})$ (Bold letters are variable.)

$= \sum_{Smokes,Cold,LungDisease} \alpha P(\mathbf{Smokes})P(\mathbf{Cold})P(\mathbf{LungDisease|Smokes})P(ChestPain = T|\mathbf{LungDisease})P(Cough = F|\mathbf{Cold, LungDisease})$

α(0.2 × 0.02 × 0.1 × 0.2 × 0.25 (TTT)

+0.2 × 0.02 × 0.9 × 0.01 × 0.5 (TTF)

+0.2 × 0.98 × 0.1 × 0.2 × 0.5 (TFT)

+0.2 × 0.98 × 0.9 × 0.01 × 0.99 (TFF)

+0.8 × 0.02 × 0.001 × 0.2 × 0.25 (FTT)

+0.8 × 0.02 × 0.999 × 0.01 × 0.5 (FTF)

+0.8 × 0.98 × 0.001 × 0.2 × 0.5 (FFT)

+0.8 × 0.98 × 0.999 × 0.01 × 0.99) (FFF)

$= \alpha \times 0.0116573184$

In [6]:
```
# 확률 계산

a = [0.2,0.2,0.2,0.2,0.8,0.8,0.8,0.8]
b = [0.02,0.02,0.98,0.98,0.02,0.02,0.98,0.98]
c = [0.1,0.9,0.1,0.9,0.001,0.999,0.001,0.999]
d = [0.2,0.8,0.2,0.8,0.2,0.8,0.2,0.8]
e = [0.25,0.5,0.5,0.99,0.25,0.5,0.5,0.99]
sol2 = 0
for i in range(8):
    sol2 += a[i]*b[i]*c[i]*d[i]*e[i]

print(f'{sol2:.10f}')
```

```
0.0116573184
```

In [7]:
```
# normalize
alpha = 1/(sol1+sol2)

prob = alpha * sol1
print(prob)
```

```
0.16446972477064223
```

$\therefore P(Cough = T|ChestPain = T) \approx 0.1645$

6.

Write a code for calculating the conditional probability $P(LungDisease = T|Smokes = T, Cough = T)$ (Attach the code and result)

Sol)

In [9]:
```
# 조건부확률 계산
predictions = model.predict_proba({
    "Smokes": "smokesTrue",
    "Cough": "coughTrue"
})

# 결과 출력
for node, prediction in zip(model.states, predictions):
    if not isinstance(prediction, str):
        for value, probability in prediction.parameters[0].items():
            if value == 'lungTrue':
                print(f'{value} : {probability:.4f}')
```

```
lungTrue : 0.7392
```

$\therefore P(LungDisease = T|Smokes = T, Cough = T) = 0.7392$

7.

Write the equation and explanation for calculating the probability of the Problem 6. (You can use the calculator)

Sol)

$P(LungDisease = T|Smokes = T, Cough = T)$

$= \sum_{Cold} \alpha P(LungDisease = T, Smokes = T, Cough = T, \mathbf{Cold})$

$= \sum_{Cold} \alpha P(Smokes = T)P(LungDisease = T|Smokes = T)P(\mathbf{Cold})P(Cough = T|\mathbf{Cold}, LungDisease = T)$

$= \alpha \times 0.2 \times 0.1 \times (0.02 \times 0.75 + 0.98 \times 0.5)$

$= \alpha \times 0.0101$

Find $P(LungDisease = F|Smokes = T, Cough = T)$ in the same way.

$P(LungDisease = F|Smokes = T, Cough = T)$

$= \sum_{Cold} \alpha P(LungDisease = F, Smokes = T, Cough = T, \mathbf{Cold})$

$= \sum_{Cold} \alpha P(Smokes = T)P(LungDisease = F|Smokes = T)P(\mathbf{Cold})P(Cough = T|\mathbf{Cold}, LungDisease = F)$

$= \alpha \times 0.2 \times 0.9 \times (0.02 \times 0.5 + 0.98 \times 0.01)$

$= \alpha \times 0.0003564$

In [10]:
```
#normalize

alpha = 1/(0.0101+0.0003564)

prob = alpha * 0.0101
print(prob)
```

```
0.7391668618266979
```

$\therefore P(LungDisease = T|Smokes = T, Cough = T) \approx 0.7392$

8.

My presentation youtube link is https://youtu.be/k_jEqfjHH1I

In [ ]:
```

```