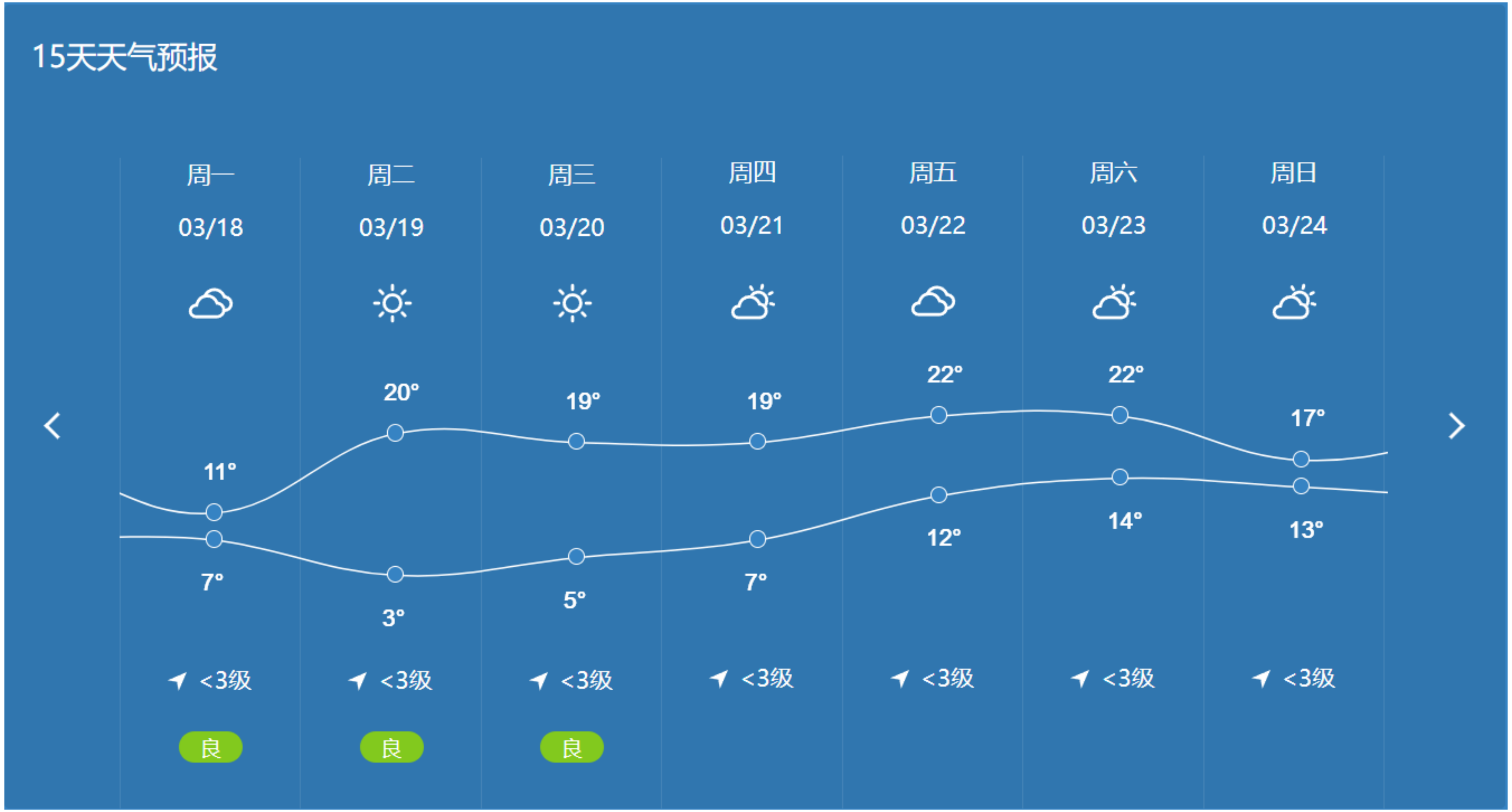
第五章

概率推理

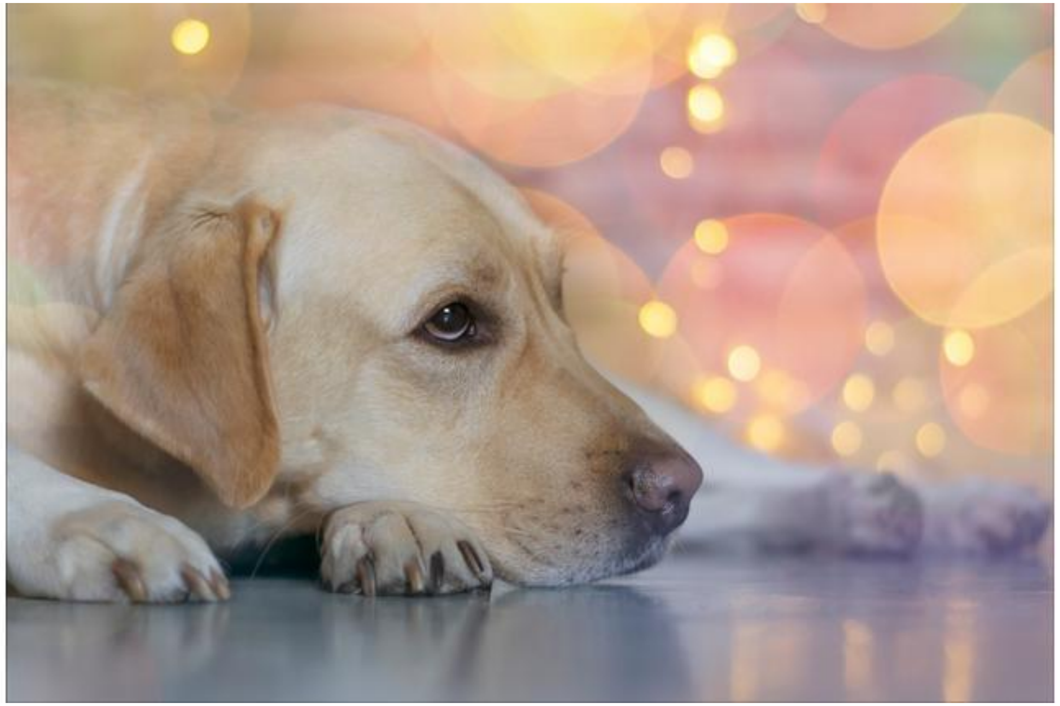
不确定性 (Uncertainty)

- **天气预报**：根据过去的天气，推理未来的天气



不确定性 (Uncertainty)

- **图像分类问题**



- 狗：95%
- 猫：3%
- 鼠：1%
- 猪：1%

概率的基本性质 (对应教材3.7.1节)

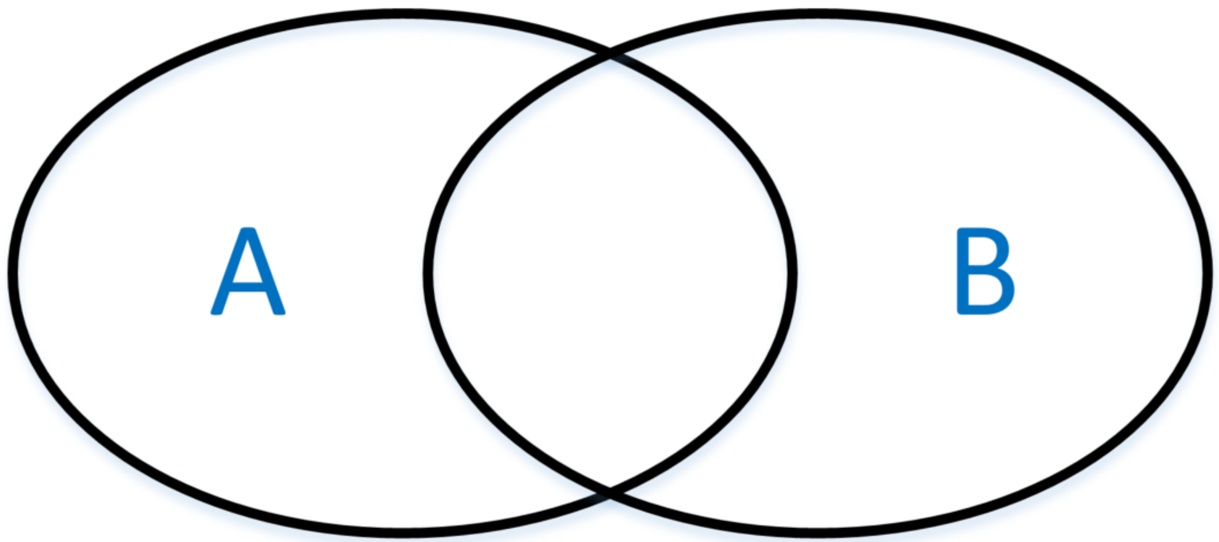
湖北大学 HUBEI UNIVERSITY

- **随机事件：**可能发生也可能不发生的实验结果，可以用大写字母A，B，C等表示
- **事件A发生的概率：**$P(A)$

$$0 \leq P(A) \leq 1$$

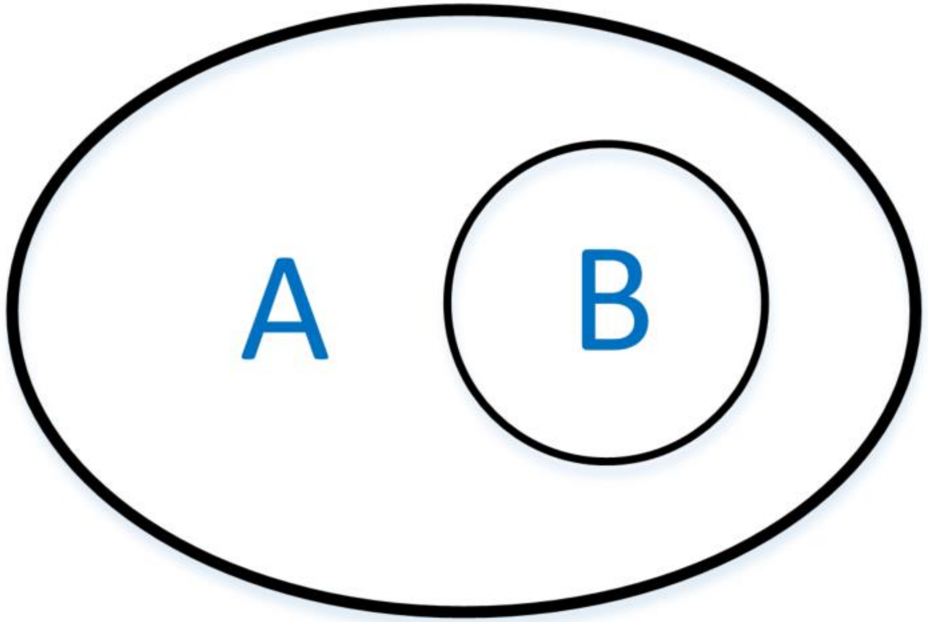
$$P(\bar{A}) = 1 - P(A)$$

- 韦恩图(Venn diagram)：



$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

- **韦恩图(Venn diagram)：**



$$P(A/B) = P(A) - P(B)$$

概率的基本性质

- 不可能事件是否等价于概率为0的事件?

- 事件A：在[0,1]任取一个实数，恰好取到0.5

$$P(A)=0$$

但A不是不可能事件

概率的基本性质

- **例：丢1个骰子**

$$\sum_{i} P(\text{点数} = i) = 1$$

- **例：丢2个骰子**

$$P(\text{点数和}=12)=\frac{1}{36}$$

$$P(\text{点数和}=7)=\frac{1}{6}$$

- 不一定所有事件都是等可能事件

- **非条件概率**：没有其他事件发生的情况下

- **条件概率：**某些事件已经发生的情况下

$$P(A\mid B)$$

$$P(\text{今天会下雨}\mid\text{昨天下雨了})$$

$$P(\text{选择路线}A\mid\text{交通状况差})$$

$$P(\text{生病}\mid\text{检查结果为阳性})$$

概率的基本性质

- **条件概率：**某些事件已经发生的情况下

$$P(A \mid B) = \frac{P(A \wedge B)}{P(B)}$$

概率的基本性质

- **条件概率：**丢2个骰子

$$P(\text{点数和}=12 \wedge \text{骰子A的点数}=6)=\frac{1}{36}$$

$$P(\text{骰子A的点数}=6)=\frac{1}{6}$$

$$P(\text{点数和}=12 \mid \text{骰子A的点数}=6)=\frac{1}{6}$$

概率的基本性质

- **条件概率**：某些事件已经发生的情况下（对应教材上的公式3.12）

$$P(A \mid B) = \frac{P(A \wedge B)}{P(B)}$$

$$P(A \wedge B) = P(B)P(A \mid B)$$

$$P(A \wedge B) = P(A)P(B \mid A)$$

湖北大学 HUBEI UNIVERSITY

- **随机变量 (random variable)**：随机试验的各种结果
  - ✓ 丢骰子：$\{1, 2, 3, 4, 5, 6\}$
  - ✓ 天气：{晴朗，多云，降雨，降雪}
  - ✓ 交通：{通畅，拥堵，严重拥堵}
  - ✓ 航班：{正点，晚点，取消}

- **概率分布 (probability distribution)：**

✓ $P(航班 = 正点) = 0.7$

✓ $P(航班 = 晚点) = 0.2$

✓ $P(航班 = 取消) = 0.1$

✓ $P(航班) = \langle 0.7, 0.2, 0.1 \rangle$

湖北大学 HUBEI UNIVERSITY

- **独立性 (independence)**：某些事件的发生不会影响另一个事件发生的概率

$$P(A \wedge B) = P(A)P(B \mid A)$$

✓ A和B相互独立时:

$$P(A \wedge B) = P(A)P(\mathrm{B})$$

湖北大学 HUBEI UNIVERSITY

- **独立性 (independence)**：某些事件的发生不会影响另一个事件发生的概率

例：

$$P(\text{骰子}A=6 \wedge \text{骰子}B=6)$$

$$=P(\text{骰子}A=6)P(\text{骰子}B=6)$$

湖北大学 HUBEI UNIVERSITY

- **独立性 (independence)**：某些事件的发生不会影响另一个事件发生的概率

例：

$$P(\text{骰子}A = 6 \wedge \text{骰子}A = 4)$$

$$\neq P(\text{骰子}A = 6)P(\text{骰子}A = 4)$$

$$= \frac{1}{6} \times \frac{1}{6} = \frac{1}{36}$$

概率的基本性质

- **独立性 (independence)**：某些事件的发生不会影响另一个事件发生的概率

例：

$$P(\text{骰子}A=6\wedge\text{骰子}A=4)$$

$$=P(\text{骰子}A=6)P(\text{骰子}A=4\,|\,\text{骰子}A=6)$$

$$=\frac{1}{6}\times 0=0$$

贝叶斯公式 (Bayes' Rule)

$$P(A \wedge B) = P(B)P(A \mid B)$$

$$P(A \wedge B) = P(A)P(B \mid A)$$

$$P(A)P(B \mid A) = P(B)P(A \mid B)$$

贝叶斯公式 (Bayes' Rule)

$$P(A)P(B \mid A) = P(B)P(A \mid B)$$

$$P(B \mid A) = \frac{P(B)P(A \mid B)}{P(A)}$$

对应教材上的公式 (3.15)

贝叶斯公式 (Bayes' Rule)

- **例**： 已知早上多云，问下午下雨的概率是多少?

✓ 下午下雨的天数中，80%的天数早上多云;

✓ 40%的天数早上多云;

✓ 10%的天数下午下雨

贝叶斯公式 (Bayes' Rule)

$$P(\text{下雨} \mid \text{多云}) = \frac{P(\text{下雨})P(\text{多云} \mid \text{下雨})}{P(\text{多云})}$$

$$= \frac{0.1 \times 0.8}{0.4}$$

$$= 0.2$$

贝叶斯公式 (Bayes' Rule)

- **意义**：根据容易获取的概率，计算不容易获取的概率

✓ **已知：**

$$P(\text{咳嗽} \mid \text{感冒})$$

✓ **可以计算：**

$$P(\text{感冒} \mid \text{咳嗽})$$

贝叶斯公式 (Bayes' Rule)

- **意义**：根据容易获取的概率，计算不容易获取的概率

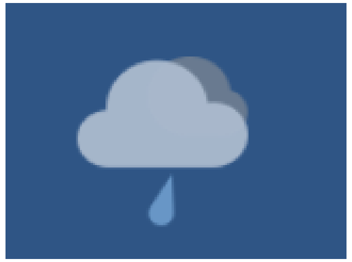
✓ **已知：**

$$P(\text{测试结果为阳性} \mid \text{疾病})$$

✓ **可以计算：**

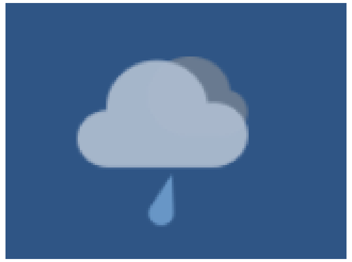
$$P(\text{疾病} \mid \text{测试结果为阳性})$$

联合概率



$C = cloud$	$C = \sim cloud$
0.4	0.6

$R = rain$	$R = \sim rain$
0.1	0.9



	$R = rain$	$R = \sim rain$
$C = cloud$	0.08	0.32
$C = \sim cloud$	0.02	0.58

联合概率

$$P(C \mid rain) = \frac{P(C \wedge rain)}{P(rain)} = \alpha P(C \wedge rain)$$

$$= \alpha\langle 0.08, 0.02\rangle = \langle 0.8, 0.2\rangle$$

	$R = rain$	$R = \sim rain$
$C = cloud$	0.08	0.32
$C = \sim cloud$	0.02	0.58

联合概率

	$R = rain$	$R = \sim rain$
$C = cloud$	0.08	0.32
$C = \sim cloud$	0.02	0.58

$$P(C = cloud)$$

$$= P(C = cloud \land R = rain) + P(C = cloud \land R = \sim rain)$$

$$= 0.08 + 0.32$$

$$= 0.40$$

概率的基本性质

全概率公式： 对应教材上的公式(3.14)

$$P(A) = P(A \wedge B) + P(A \wedge \sim B)$$

$$P(X = x_i) = \sum_j P\big(X = x_i \wedge Y = y_j\big)$$

概率的基本性质

$$P(A) = P(A \wedge B) + P(A \wedge \sim B)$$

$$P(A) = P(A \mid B)P(B) + P(A \mid \sim B)P(B)$$

$$P(X = x_i) = \sum_j P\big(X = x_i \mid Y = y_j\big)P(Y = y_j)$$

概率推理方法

- **基于知识的推理：**根据已知信息推理未知信息，使用等价规则求得合取范式，再使用消解推理规则推出矛盾
- **概率推理：**条件概率，使用贝叶斯公式
- **教材上的例3.10：**

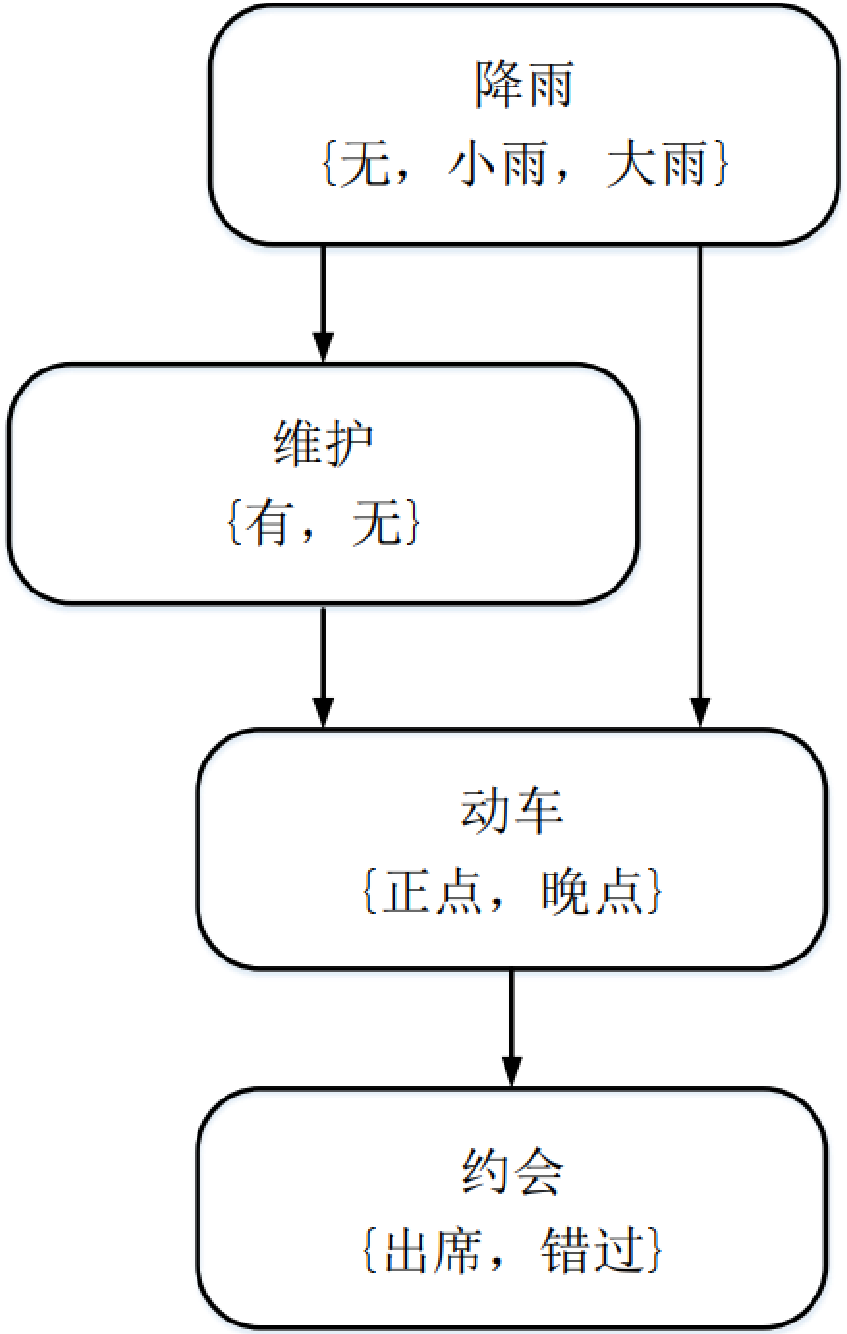
$$P(H_1 \mid E) = \frac{P(H_1)P(E \mid H_1)}{P(E)} \quad \text{(贝叶斯公式)}$$

$$= \frac{P(H_1)P(E \mid H_1)}{\sum_{i=1}^{3} P(E \mid H_i)P(H_i)} \quad \text{(对分母使用全概率公式)}$$

贝叶斯网络（Bayesian network）

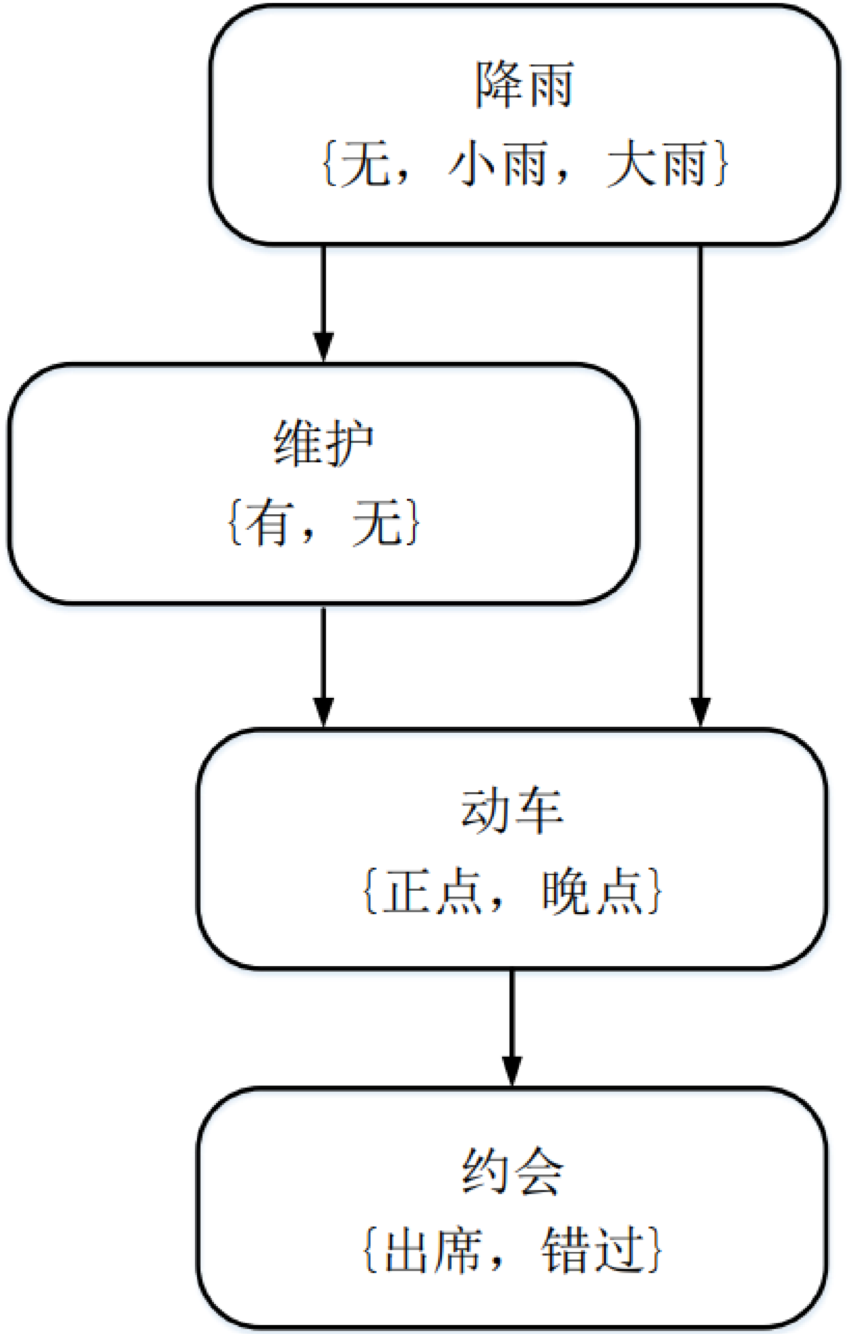
- **定义：** 使用有向无环图表示随机变量之间的依赖关系
  - ✓ 每个结点代表一个随机变量
  - ✓ 从 X 到 Y 的有向边代表 X 是 Y 的一个父结点
  - ✓ 每个结点 X 都有一个概率分布$P(X \mid Parent(X))$

贝叶斯网络（Bayesian network）



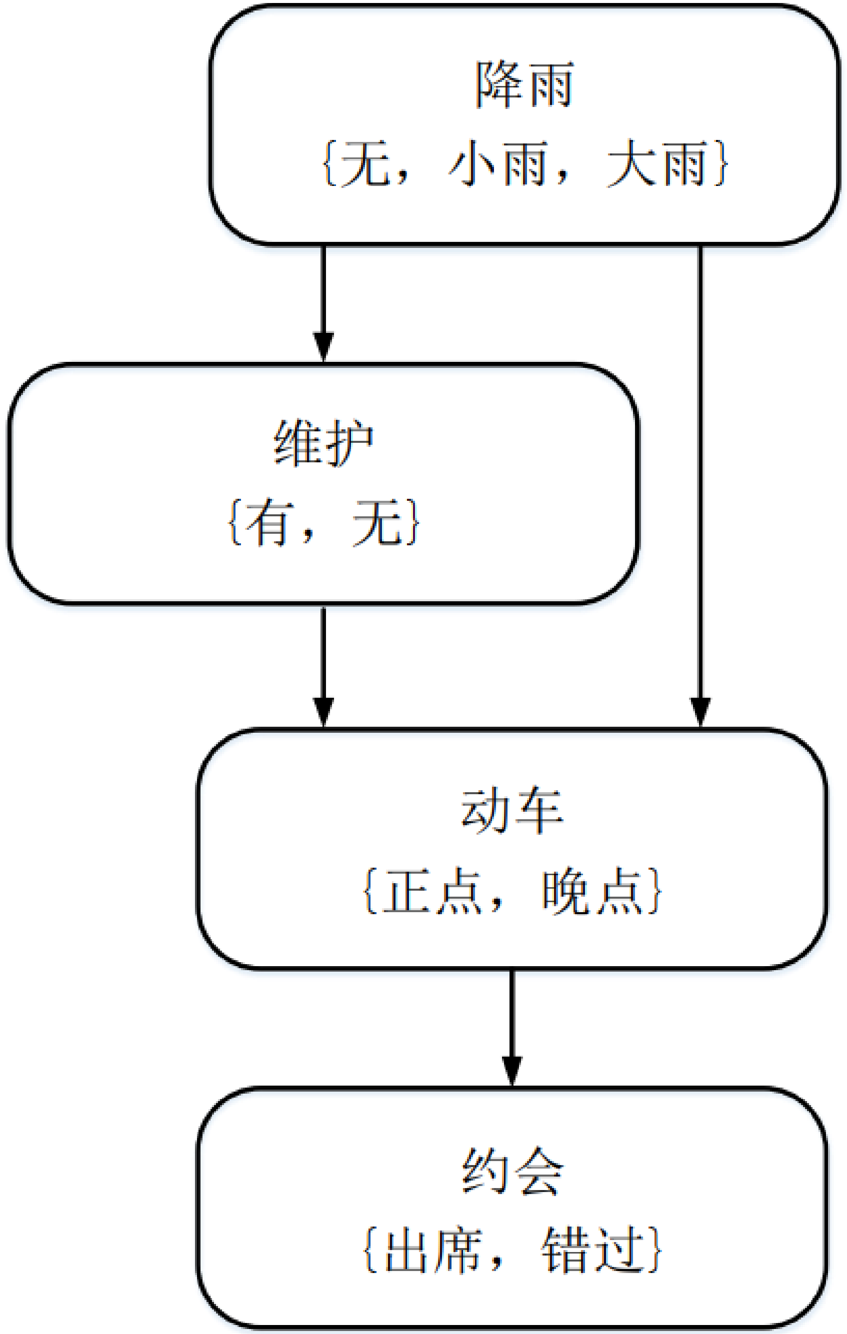
无	小雨	大雨
0.7	0.2	0.1

贝叶斯网络（Bayesian network）



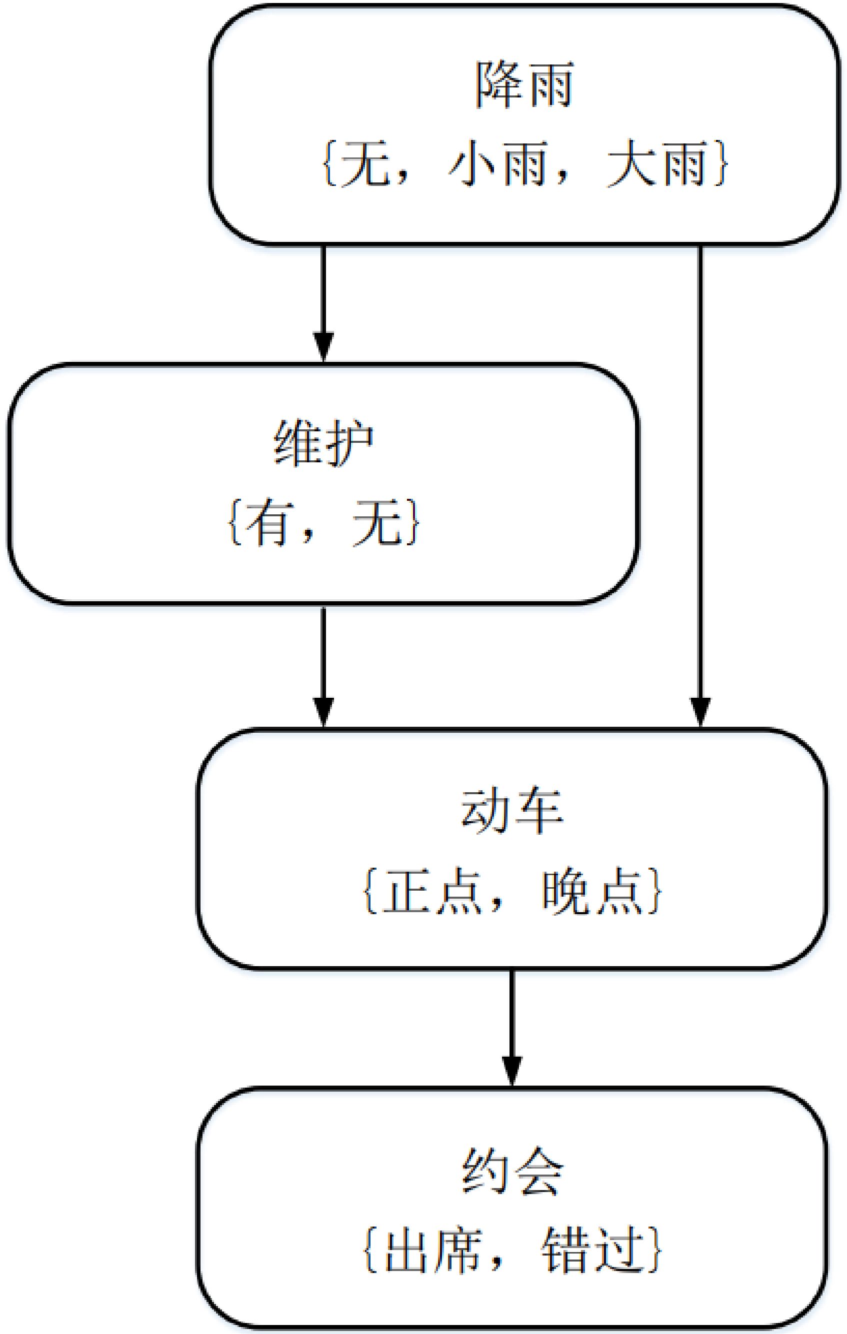
降雨	有维护	无维护
无	0.4	0.6
小雨	0.2	0.8
大雨	0.1	0.9

贝叶斯网络（Bayesian network）



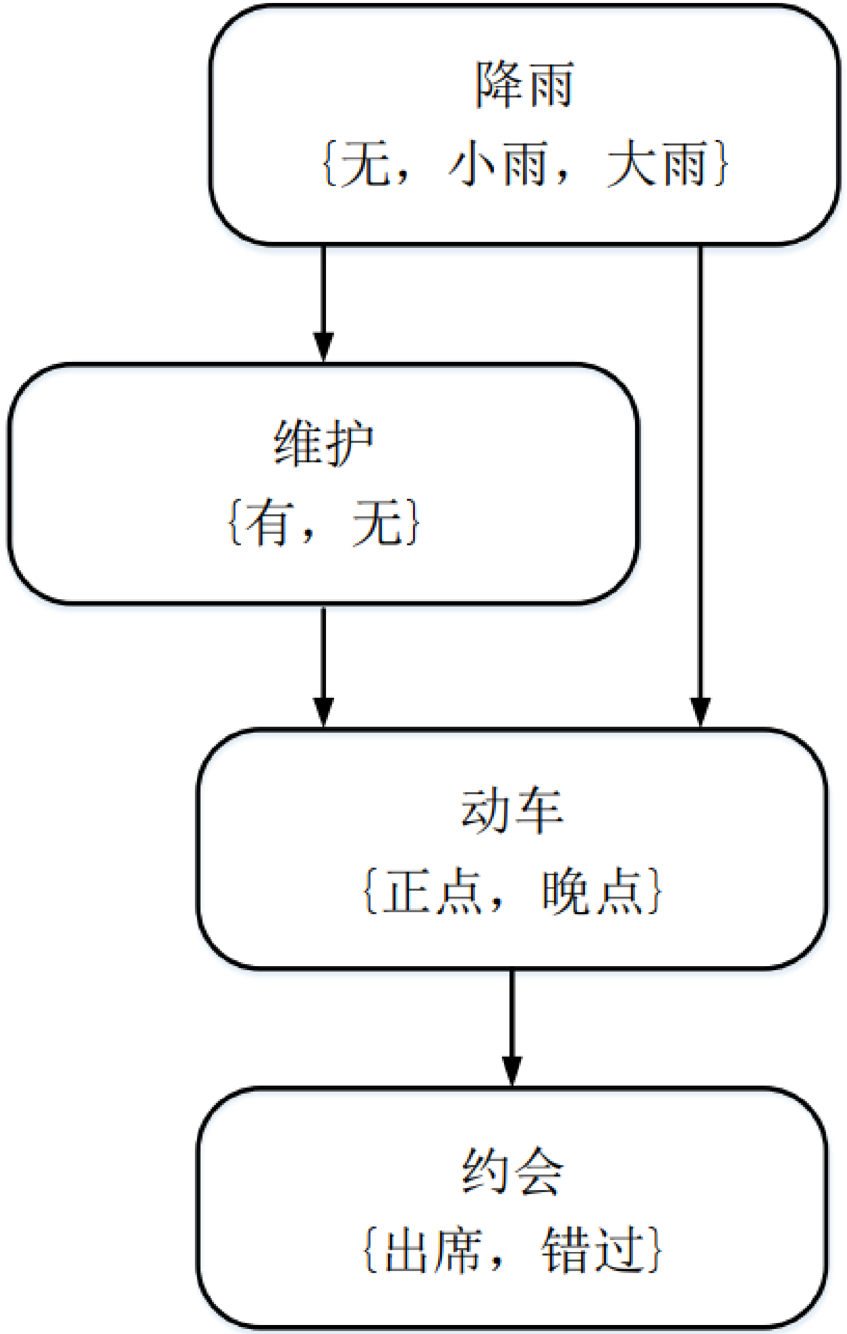
降雨	维护	动车正点	动车晚点
无	有	0.8	0.2
无	无	0.9	0.1
小雨	有	0.6	0.4
小雨	无	0.7	0.3
大雨	有	0.4	0.6
大雨	无	0.5	0.5

贝叶斯网络（Bayesian network）



动车	出席约会	错过约会
正点	0.9	0.1
晚点	0.6	0.4

贝叶斯网络（Bayesian network）

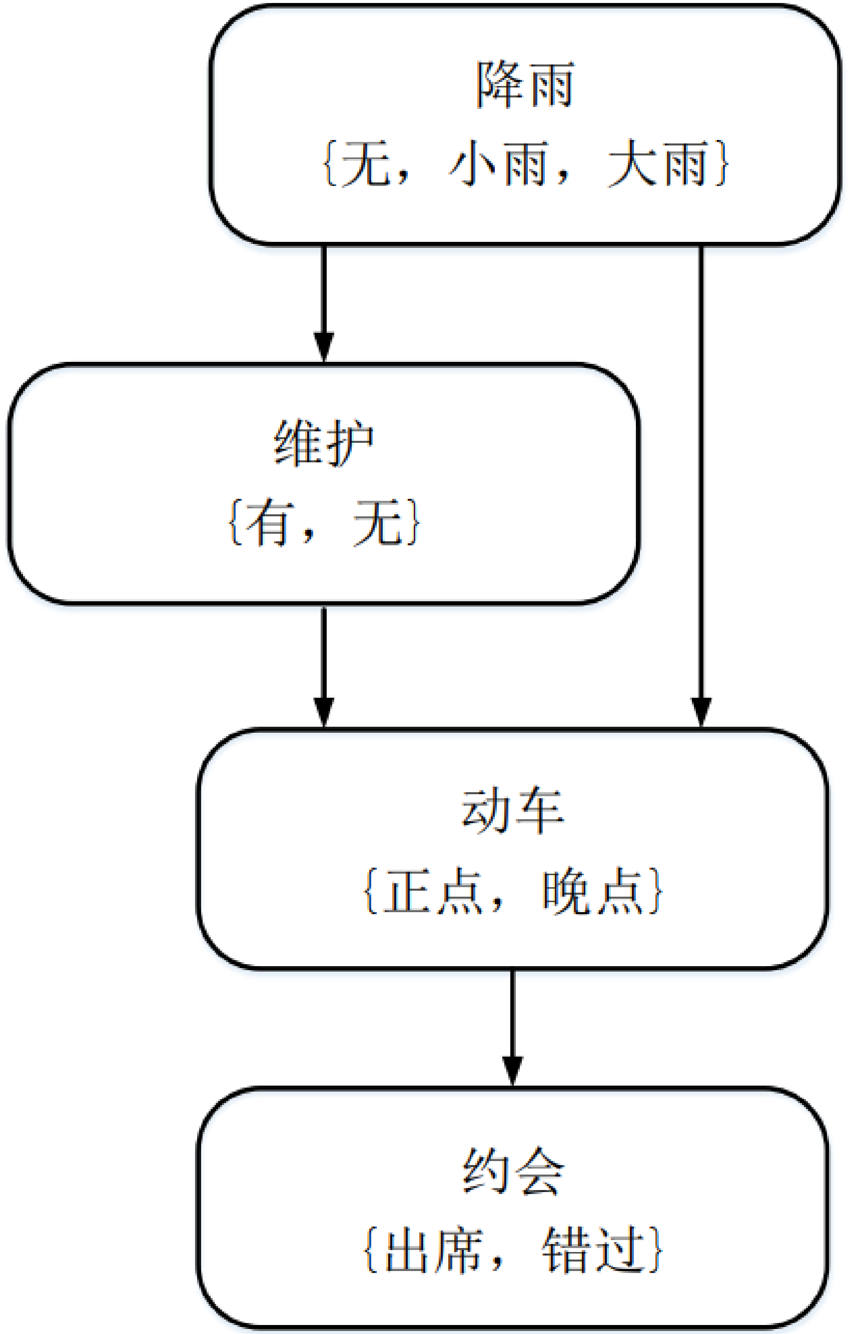


- **计算联合概率**

$$P(\text{小雨，无维护})$$

$$= P(\text{小雨})P(\text{无维护} \mid \text{小雨})$$

贝叶斯网络（Bayesian network）

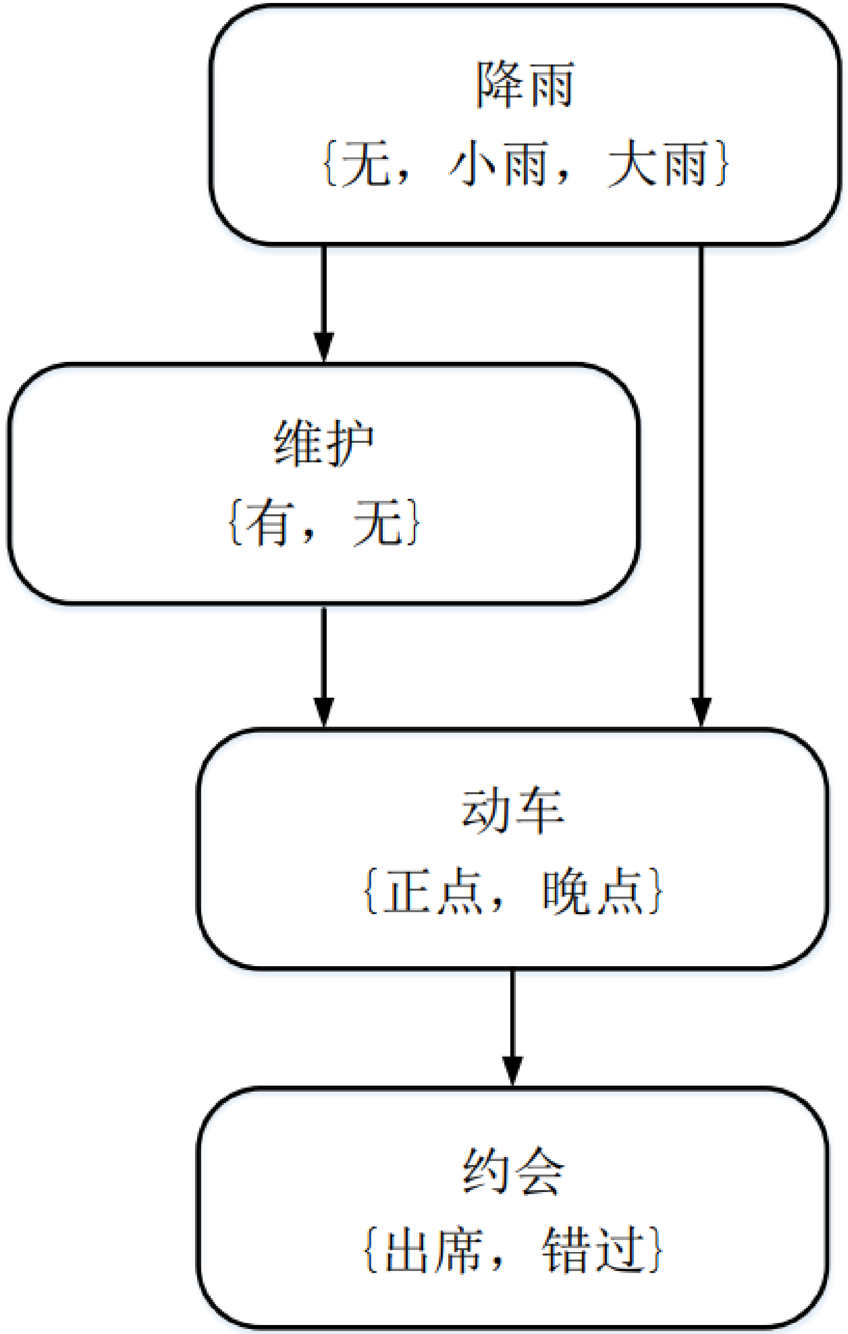


- **计算联合概率**

P(小雨，无维护，动车晚点)

$= P$(小雨)P(无维护 | 小雨)P(动车晚点 |小雨，无维护)

贝叶斯网络（Bayesian network）



- **计算联合概率**

$$P(\text{小雨，无维护，动车晚点，错过约会})$$
$$= P(\text{小雨})P(\text{无维护} \mid \text{小雨})P(\text{动车晚点} \mid \text{小雨，无维护})$$
$$P(\text{错过约会} \mid \text{动车晚点})$$

基于贝叶斯网络的推理

- **查询变量** X：需要计算概率分布的变量
- **证据变量 (evidence variable)** E：在事件 e 中观察到的变量
- **隐藏变量 (hidden variable)** Y: 既不是查询变量，也不是证据变量

✓目标：计算 $P(X \mid e)$

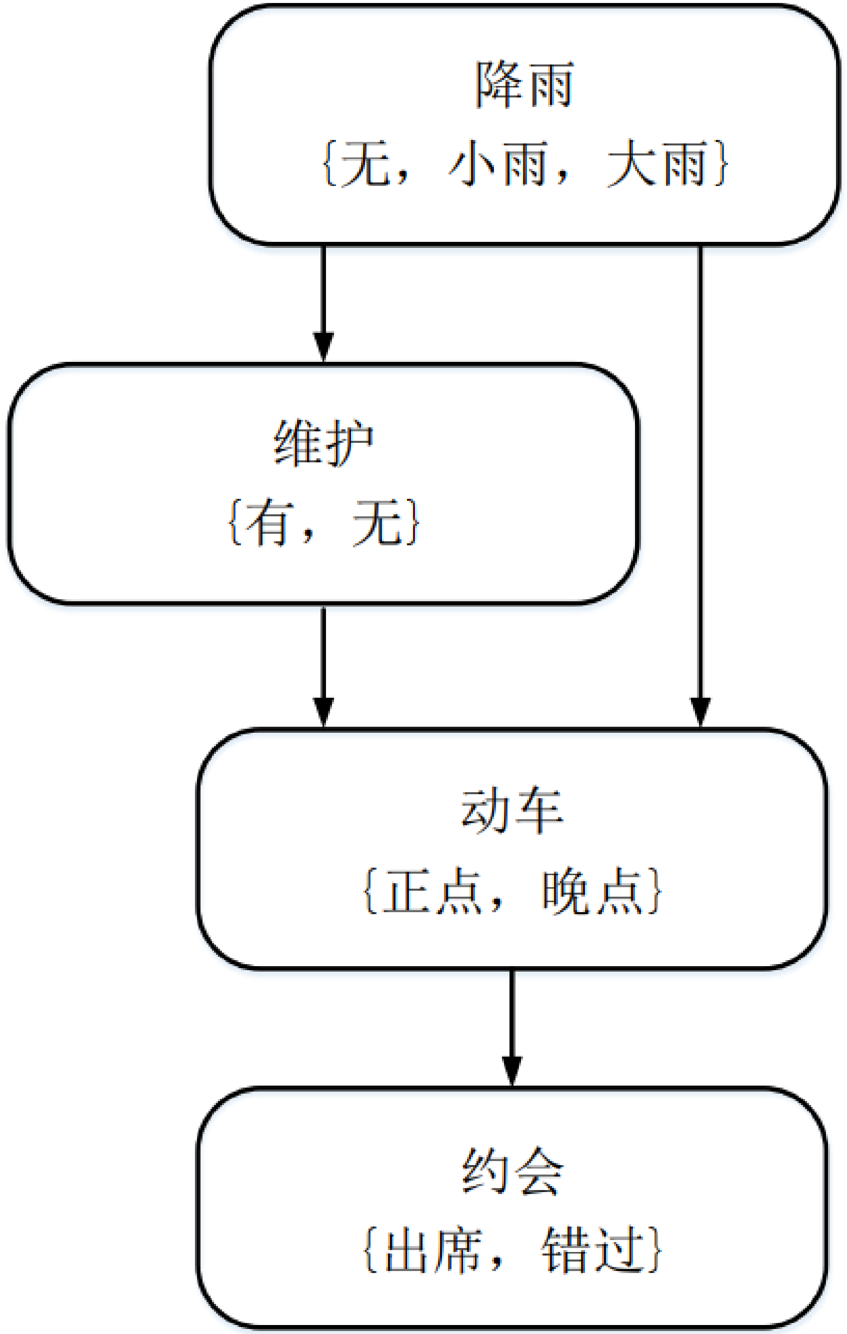
基于贝叶斯网络的推理

$$P(\text{约会} \mid \text{小雨，无维护})$$

$$= \alpha P(\text{约会，小雨，无维护})$$

$$= \alpha [P(\text{约会，小雨，无维护，动车正点})$$

$$+ P(\text{约会，小雨，无维护，动车晚点})]$$

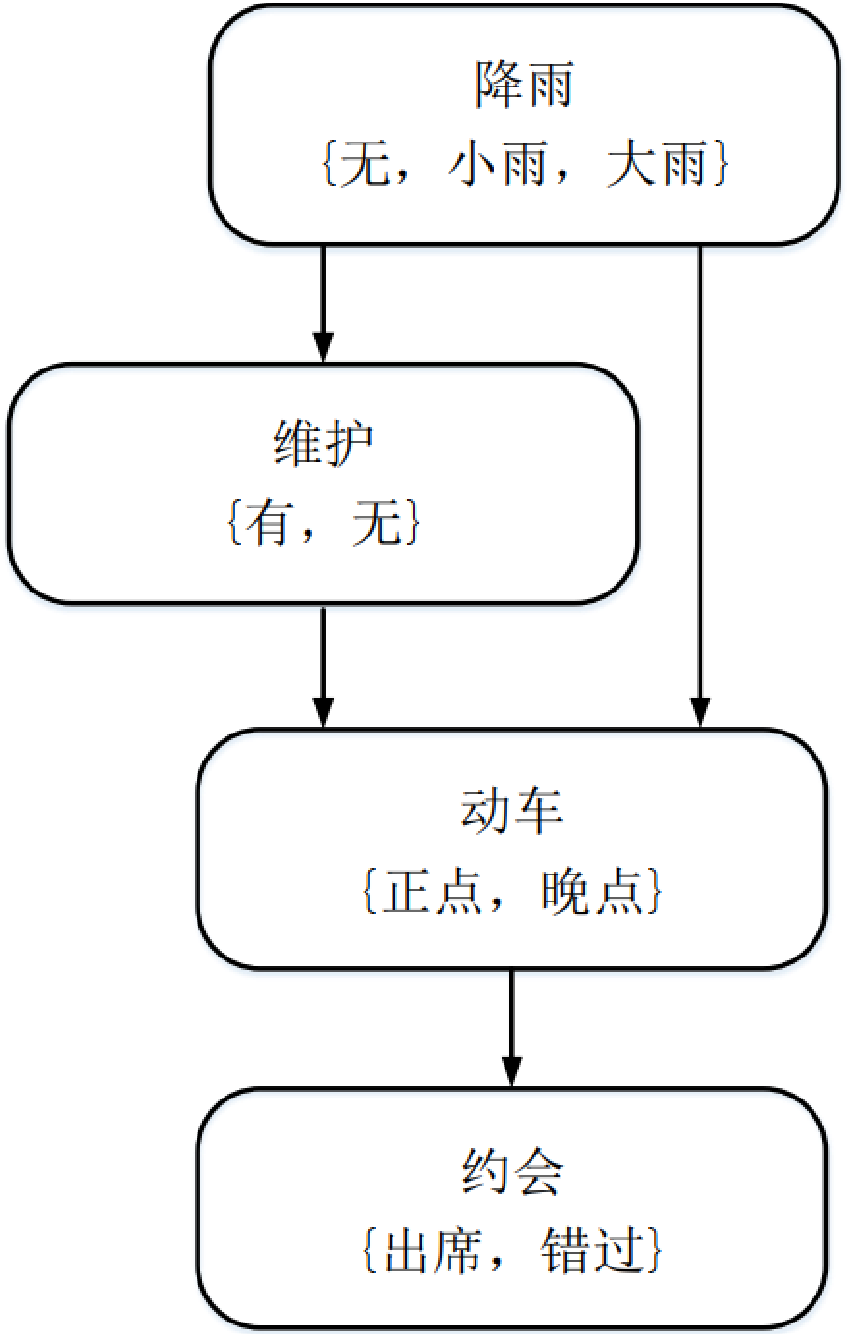


基于贝叶斯网络的推理

$$P(X \mid e) = \alpha P(X, e)$$

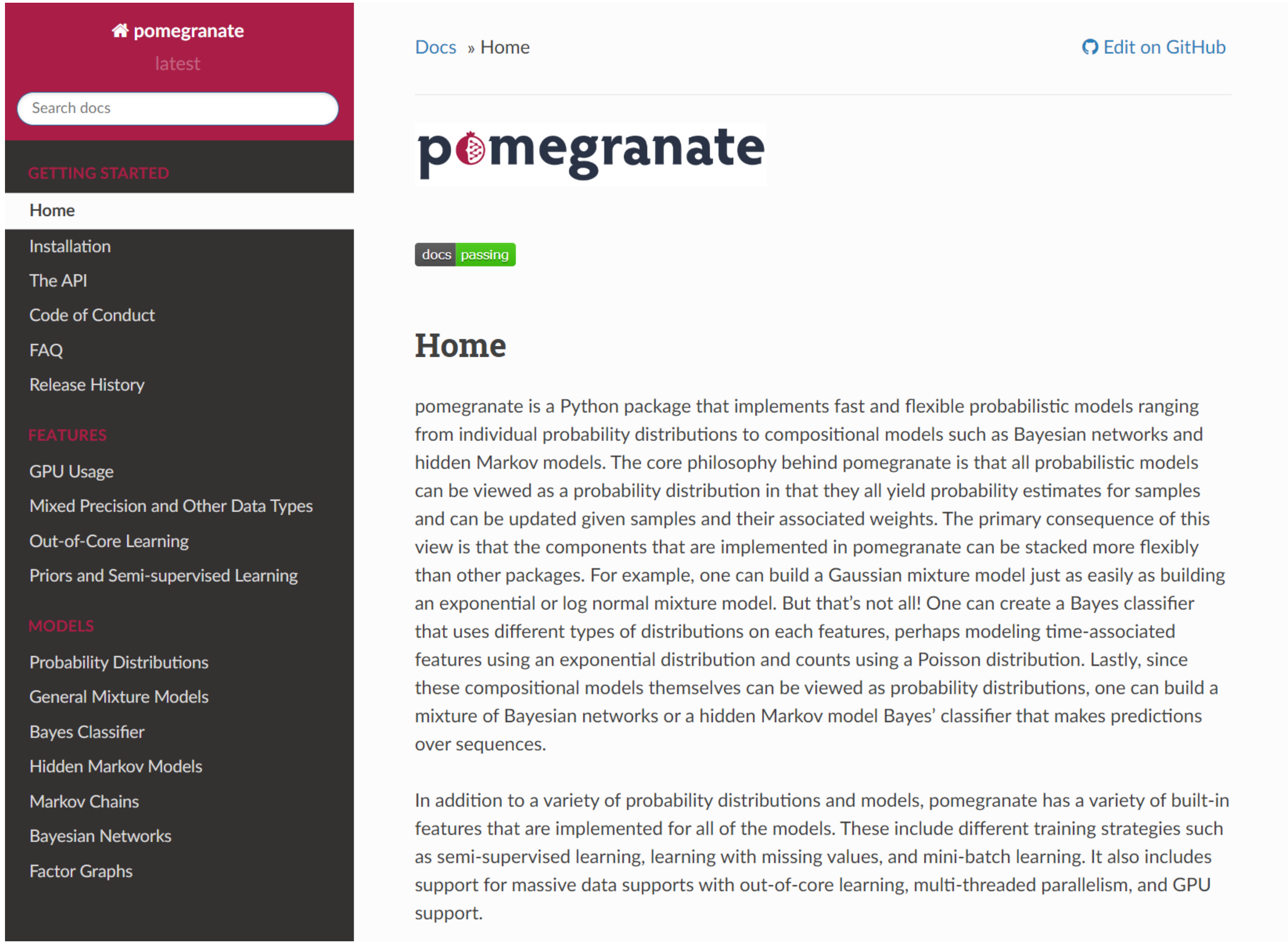
$$= \alpha \sum_{y} \mathrm{P}(X, e, y)$$

- X：需要查询的变量
- e：已经观察到的证据
- y：隐藏变量所有可能的取值



Pomegranate

- 用于实现概率模型（如贝叶斯网络，隐马尔科夫模型）的python库
- 官网：https://pomegranate.readthedocs.io/en/latest/index.html



pomegranate
latest
Search docs

GETTING STARTED

- Home
- Installation
- The API
- Code of Conduct
- FAQ
- Release History

FEATURES

- GPU Usage
- Mixed Precision and Other Data Types
- Out-of-Core Learning
- Priors and Semi-supervised Learning

MODELS

- Probability Distributions
- General Mixture Models
- Bayes Classifier
- Hidden Markov Models
- Markov Chains
- Bayesian Networks
- Factor Graphs

Docs » Home
Edit on GitHub

pomegranate

docs passing

Home

pomegranate is a Python package that implements fast and flexible probabilistic models ranging from individual probability distributions to compositional models such as Bayesian networks and hidden Markov models. The core philosophy behind pomegranate is that all probabilistic models can be viewed as a probability distribution in that they all yield probability estimates for samples and can be updated given samples and their associated weights. The primary consequence of this view is that the components that are implemented in pomegranate can be stacked more flexibly than other packages. For example, one can build a Gaussian mixture model just as easily as building an exponential or log normal mixture model. But that's not all! One can create a Bayes classifier that uses different types of distributions on each features, perhaps modeling time-associated features using an exponential distribution and counts using a Poisson distribution. Lastly, since these compositional models themselves can be viewed as probability distributions, one can build a mixture of Bayesian networks or a hidden Markov model Bayes' classifier that makes predictions over sequences.

In addition to a variety of probability distributions and models, pomegranate has a variety of built-in features that are implemented for all of the models. These include different training strategies such as semi-supervised learning, learning with missing values, and mini-batch learning. It also includes support for massive data supports with out-of-core learning, multi-threaded parallelism, and GPU support.

Pomegranate

```
from pomegranate import *

rain = Node(DiscreteDistribution({
    'none': 0.7,
    'light': 0.2,
    'heavy': 0.1
}), name='rain')
```

Pomegranate

```
maintainance = Node(ConditionalProbabilityTable([
    ['none', 'yes', 0.4],
    ['none', 'no', 0.6],
    ['light', 'yes', 0.2],
    ['light', 'no', 0.8],
    ['heavy', 'yes', 0.1],
    ['heavy', 'no', 0.9],
], [rain.distribution]), name='maintainance')
```

Pomegranate

```
train = Node(ConditionalProbabilityTable([
    ['none', 'yes', 'on time', 0.8],
    ['none', 'yes', 'delay', 0.2],
    ['none', 'no', 'on time', 0.9],
    ['none', 'no', 'delay', 0.1],
    ['light', 'yes', 'on time', 0.6],
    ['light', 'yes', 'delay', 0.4],
    ['light', 'no', 'on time', 0.7],
    ['light', 'no', 'delay', 0.3],
    ['heavy', 'yes', 'on time', 0.4],
    ['heavy', 'yes', 'delay', 0.6],
    ['heavy', 'no', 'on time', 0.5],
    ['heavy', 'no', 'delay', 0.5],
], [rain.distribution, maintainance.distribution]), name='train')
```

Pomegranate

```
appointment = Node(ConditionalProbabilityTable([
    ['on time', 'attend', 0.9],
    ['on time', 'miss', 0.1],
    ['delay', 'attend', 0.6],
    ['delay', 'miss', 0.4],
], [train.distribution]), name='appointment')
```

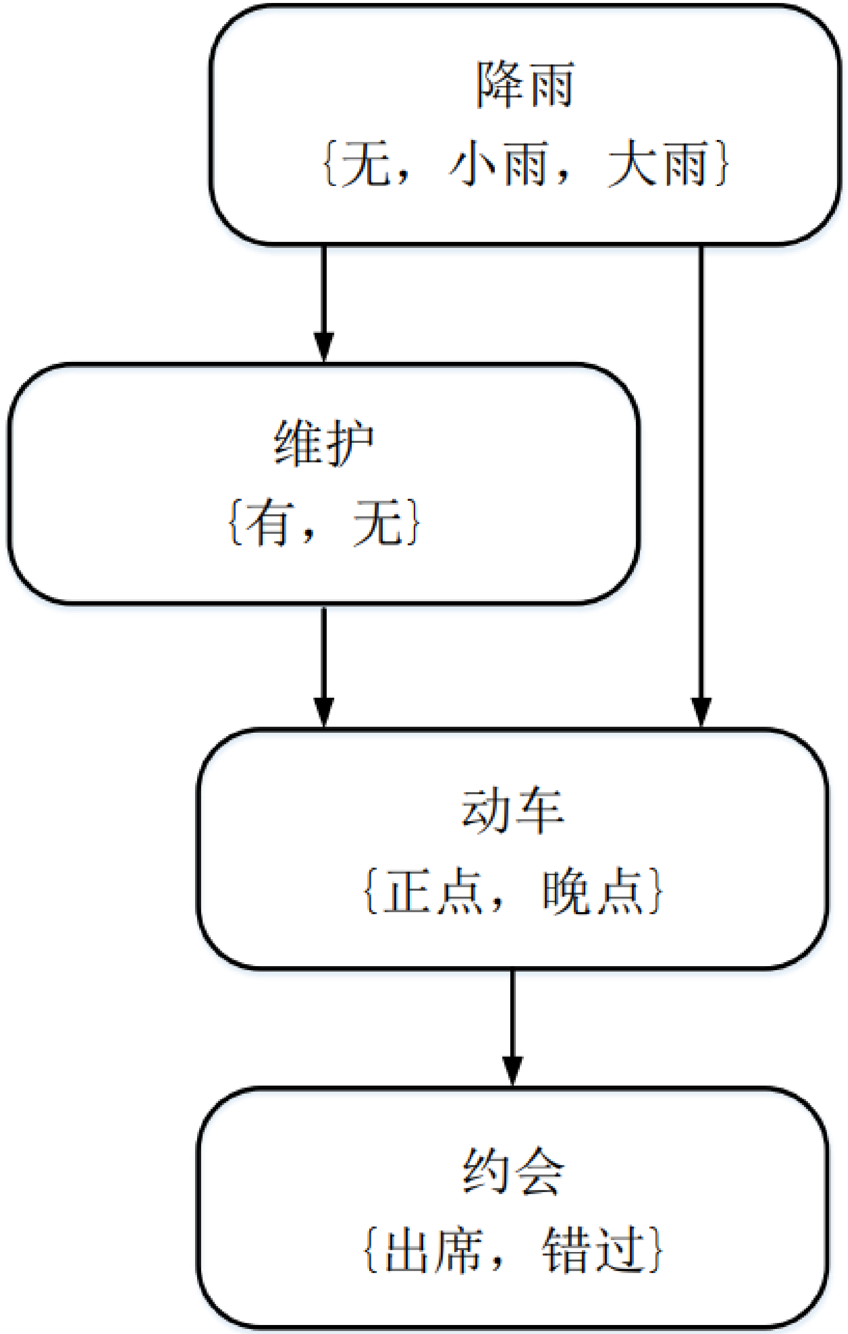
Pomegranate

```
model = BayesianNetwork()
# 添加贝叶斯网络中的结点
model.add_states(rain, maintainance, train, appointment)

# 添加贝叶斯网络中的有向边
model.add_edge(rain, maintainance)
model.add_edge(rain, train)
model.add_edge(maintainance, train)
model.add_edge(train, appointment)

model.bake()
```

贝叶斯网络（Bayesian network）



- **计算联合概率**

$$P(\text{小雨，无维护，动车晚点，错过约会})$$

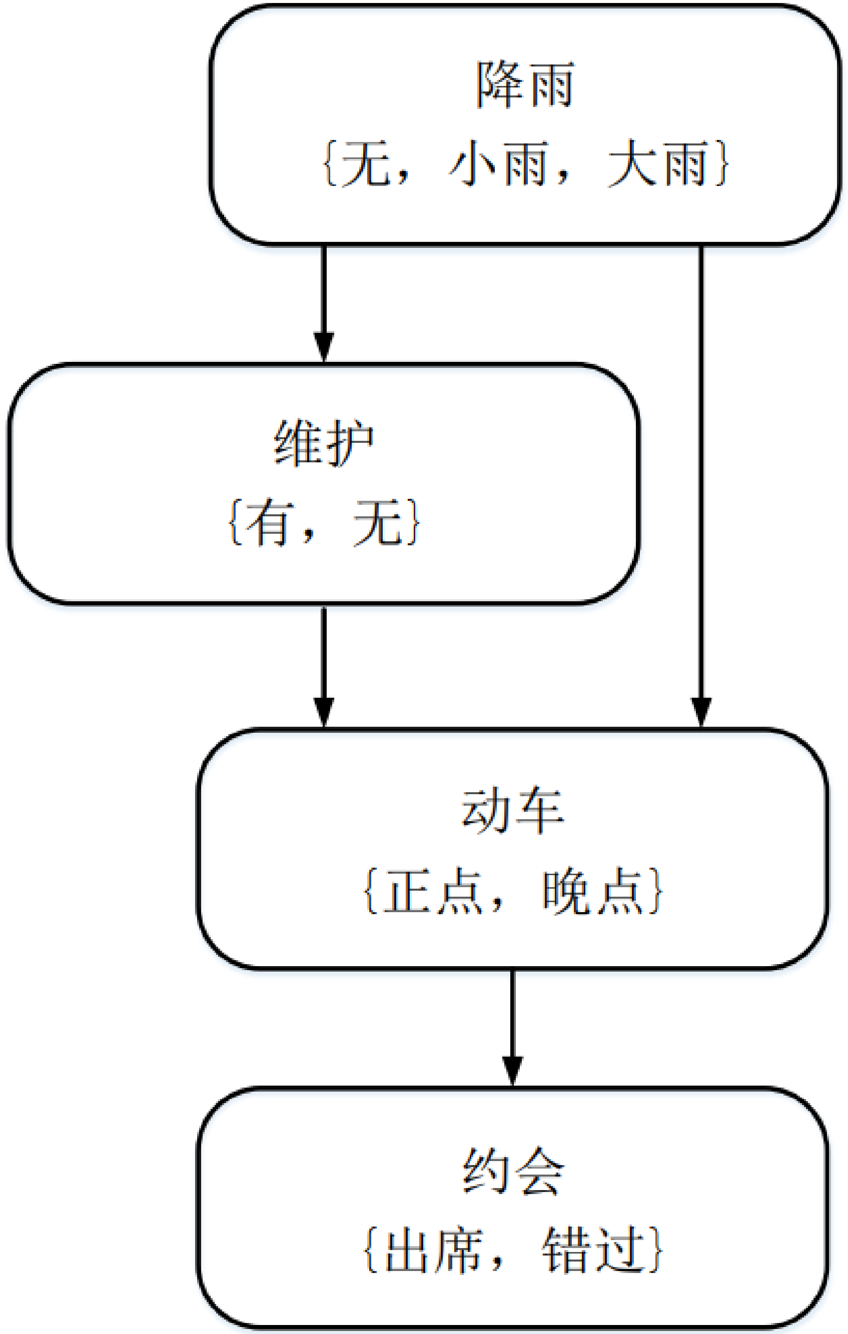
$$= P(\text{小雨})P(\text{无维护} \mid \text{小雨})P(\text{动车晚点} \mid \text{小雨，无维护})$$

$$P(\text{错过约会} \mid \text{动车晚点})$$

$$= 0.2 \times 0.8 \times 0.3 \times 0.4$$

$$= 0.0192$$

贝叶斯网络（Bayesian network)



- **计算联合概率**

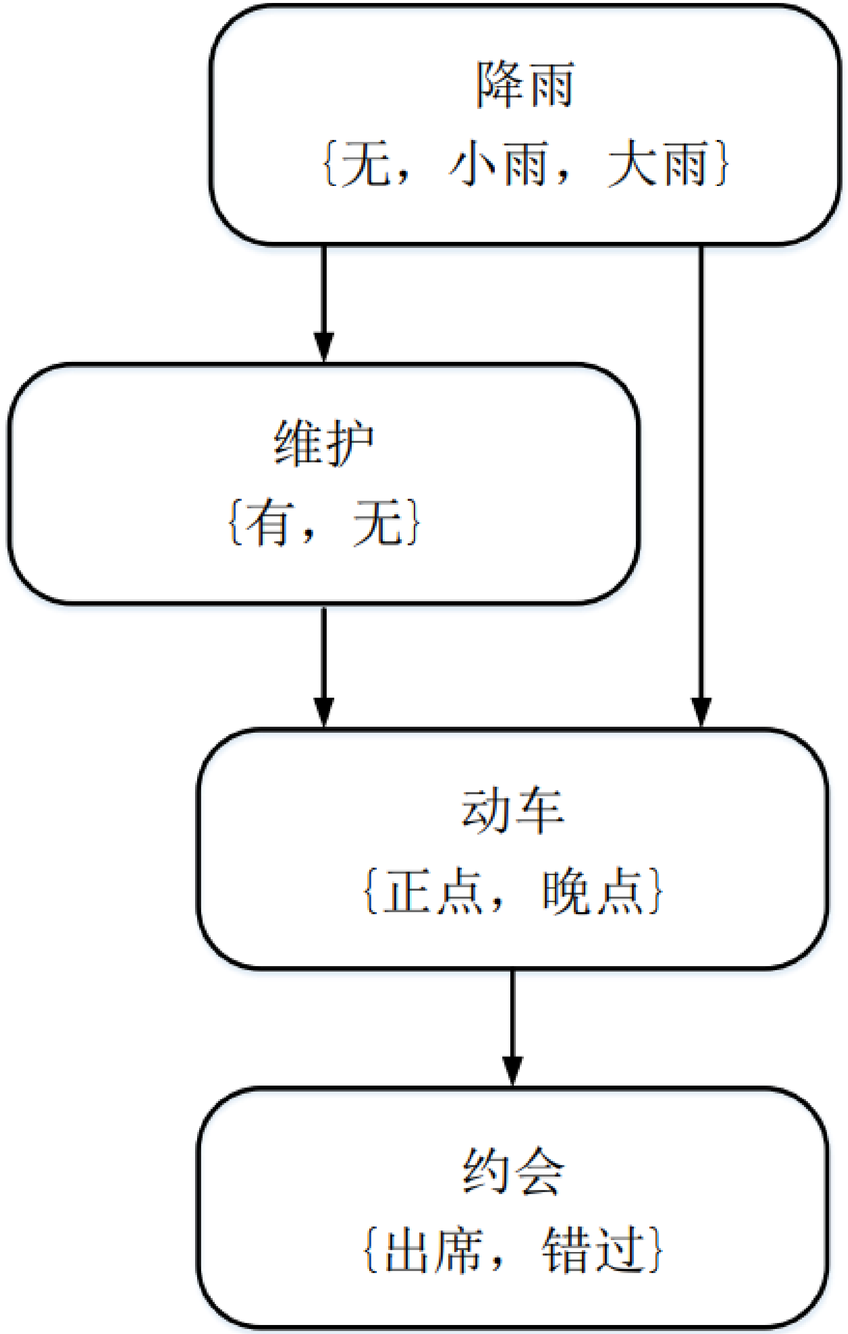
```
prob = model.probability([['light', 'no',
                           'delay', 'miss']])
print(prob)
```

- **运行结果：**

```
D:\Software\anaconda3\envs\matplot\python.exe
0.0192

Process finished with exit code 0
```

贝叶斯网络（Bayesian network)



- **计算条件概率**

P(约会 | 小雨，无维护)

$= \alpha P$(约会，小雨，无维护)

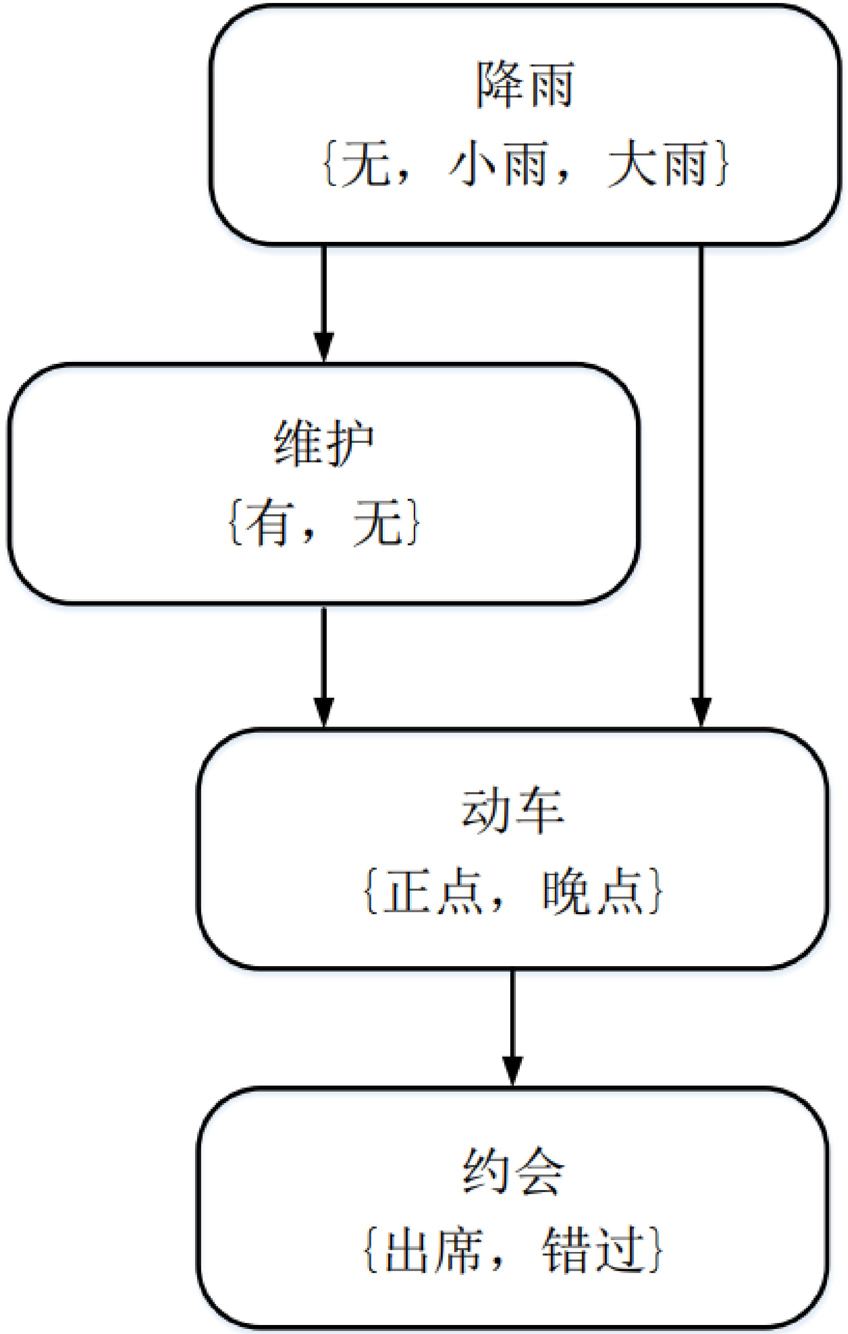
$= \alpha$[P(约会，小雨，无维护，动车正点)

+P(约会，小雨，无维护，动车晚点)]

$$= \frac{1}{0.2 \times 0.8}(0.2 \times 0.8 \times 0.7 \times 0.9 + 0.2 \times 0.8 \times 0.3 \times 0.6)$$

$= 0.81$

贝叶斯网络（Bayesian network)



- **计算条件概率**

```
predictions = model.predict_proba({
    'rain': 'light',
    'maintainance': 'no'
})
```

- **运行结果：**

```
D:\Software\anaconda3\envs\matplot\python.exe
rain: light
maintainance: no
train
on time : 0.7000
delay : 0.3000
appointment
attend : 0.8100
miss : 0.1900

Process finished with exit code 0
```

- **朴素贝叶斯分类器 (Naïve Bayes Classifier)**
  - ✓ 可以解决分类问题 (classification)
  - ✓ 假设A，B，C为特征 (feature)
  - ✓ 假设 X 为类别 (class / category / label)

湖北大学 HUBEI UNIVERSITY

训练集（已知的历史数据）：

样本 (sample)	1	2	3	4	5
瓜蒂（特征A）	脱落	未脱	未脱	脱落	脱落
形状（特征B）	圆形	尖形	圆形	尖形	圆形
颜色（特征C）	深绿	浅绿	深绿	浅绿	浅绿
类别（X）	熟瓜	生瓜	生瓜	熟瓜	熟瓜

根据训练集中的样本，判断：瓜蒂脱落，圆形，深绿的瓜是熟瓜还是生瓜

解法：分别求出以下两个条件概率

$$P(X=\text{熟瓜} \mid A=\text{脱落}, B=\text{圆形}, C=\text{深绿})$$

$$P(X=\text{生瓜} \mid A=\text{脱落}, B=\text{圆形}, C=\text{深绿})$$

根据条件概率公式：

$$P(X \mid A, B, C)$$

$$= \frac{P(X, A, B, C)}{P(A, B, C)}$$

$$= \frac{P(X)P(A \mid X)P(B \mid X, A)P(C \mid X, A, B)}{P(A, B, C)}$$

朴素贝叶斯

$$\frac{P(X)P(A\mid X)P(B\mid X,A)P(C\mid X,A,B)}{P(A,B,C)}$$

训练集：

样本 (sample)	1	2	3	4	5
瓜蒂（特征A）	脱落	未脱	未脱	脱落	脱落
形状（特征B）	圆形	尖形	圆形	尖形	圆形
颜色（特征C）	深绿	浅绿	深绿	浅绿	浅绿
类别（X）	熟瓜	生瓜	生瓜	熟瓜	熟瓜

$$P(X=\text{熟瓜})=\frac{3}{5}$$

$$P(A=\text{脱落}\mid X=\text{熟瓜})=\frac{3}{3}=1$$

$$P(B=\text{圆形}\mid X=\text{熟瓜，}A=\text{脱落})=\frac{2}{3}$$

随着特征的增加，条件概率的计算越来越麻烦

$$\frac{P(X)P(A\mid X)P(B\mid X,A)P(C\mid X,A,B)}{P(A,B,C)}$$

$$=\alpha\cdot P(X)P(A\mid X)P(B\mid X,A)P(C\mid X,A,B)$$

朴素：假设特征A，B，C之间两两独立

原条件概率公式可转化为：

$$\alpha\cdot P(X)P(A\mid X)P(B\mid X,A)P(C\mid X,A,B)$$

$$=\alpha\cdot P(X)P(A\mid X)P(B\mid X)P(C\mid X)$$

朴素贝叶斯

训练集（已知的历史数据）：

样本 (sample)	1	2	3	4	5
瓜蒂（特征A）	脱落	未脱	未脱	脱落	脱落
形状（特征B）	圆形	尖形	圆形	尖形	圆形
颜色（特征C）	深绿	浅绿	深绿	浅绿	浅绿
类别（X）	熟瓜	生瓜	生瓜	熟瓜	熟瓜

$$P(X = \text{熟瓜} \mid A = \text{脱落}, B = \text{圆形}, C = \text{深绿})$$

$$= \alpha \cdot P(X = \text{熟瓜}) P(A = \text{脱落} | X = \text{熟瓜}) P(B = \text{圆形} | X = \text{熟瓜}) P(C = \text{深绿} | X = \text{熟瓜})$$

$$= \alpha \cdot \frac{3}{5} \times \frac{3}{3} \times \frac{2}{3} \times \frac{2}{3}$$

朴素贝叶斯

训练集（已知的历史数据）：

样本 (sample)	1	2	3	4	5
瓜蒂（特征A）	脱落	未脱	未脱	脱落	脱落
形状（特征B）	圆形	尖形	圆形	尖形	圆形
颜色（特征C）	深绿	浅绿	深绿	浅绿	浅绿
类别（X）	熟瓜	生瓜	生瓜	熟瓜	熟瓜

$$P(X=\text{生瓜}\mid A=\text{脱落},B=\text{圆形},C=\text{深绿})$$

$$=\alpha\cdot P(X=\text{生瓜})P(A=\text{脱落}|X=\text{生瓜})P(B=\text{圆形}|X=\text{生瓜})P(C=\text{深绿}|X=\text{生瓜})$$

$$=\alpha\cdot\frac{2}{5}\times\frac{0}{2}\times\frac{1}{2}\times\frac{1}{2}$$

朴素贝叶斯

$$P(X=\text{熟瓜}\mid A=\text{脱落},B=\text{圆形},C=\text{深绿})$$
$$=\alpha\cdot P(X=\text{熟瓜})P(A=\text{脱落}|X=\text{熟瓜})P(B=\text{圆形}|X=\text{熟瓜})P(C=\text{深绿}|X=\text{熟瓜})$$
$$=\alpha\cdot\frac{3}{5}\times\frac{3}{3}\times\frac{2}{3}\times\frac{2}{3}$$

$$P(X=\text{生瓜}\mid A=\text{脱落},B=\text{圆形},C=\text{深绿})$$
$$=\alpha\cdot P(X=\text{生瓜})P(A=\text{脱落}|X=\text{生瓜})P(B=\text{圆形}|X=\text{生瓜})P(C=\text{深绿}|X=\text{生瓜})$$
$$=\alpha\cdot\frac{2}{5}\times\frac{0}{2}\times\frac{1}{2}\times\frac{1}{2}$$

因为$P(X=\text{生瓜}\mid A=\text{脱落},B=\text{圆形},C=\text{深绿})>P(X=\text{生瓜}\mid A=\text{脱落},B=\text{圆形},C=\text{深绿})$

所以瓜蒂脱落，圆形，深绿的瓜更可能是熟瓜

理论课 第4次平时作业 （4月9日上课前交给班长）

湖北大学 HUBEI UNIVERSITY

- **编程实现PPT中的贝叶斯网络**

  ✓ 在这个例子中，任选一个联合概率进行计算，将代码和计算结果截图（本页中的例子除外）；

  ✓ 手动验算。

例：

P(小雨，无维护，动车晚点，错过约会)

$=P$(小雨)P(无维护 | 小雨)P(动车晚点 |小雨，无维护)

P(错过约会 | 动车晚点)

$= 0.2 \times 0.8 \times 0.3 \times 0.4$

$= 0.0192$

理论课 第4次平时作业 （4月9日上课前交给班长）

- 在PPT上的例子中，使用朴素贝叶斯算法，判断瓜蒂未脱、尖形、浅色的瓜更可能是生瓜还是熟瓜。（要求写出详细过程）
- 阅读教材第3.5节，完成教材第130页的习题3-11。

结束语

谢 谢！