Harvard · CS50-Al | Introduction to Artificial Intelligence with Python (2020)

CS50-AI (2020)· 课程资料包 @ShowMeAI









视频 中英双语字幕

课件 一键打包下载

笙记 官方笔记翻译

代码 作业项目解析



视频·B站[扫码或点击链接]

https://www.bilibili.com/video/BV1AQ4y1y7wy



课件 & 代码·博客[扫码或点击链接]

http://blog.showmeai.tech/harvard-cs50-ai

深度优先 广度优先

贝叶斯

聚类 马尔可夫

A* 算法 自然语言处理

SVM RNN

强化学习 tensorflow 反向传播

监督学习

语法解析

神经网络 无监督

CNN

搜索

Awesome Al Courses Notes Cheatsheets 是 **ShowMeAI** 资料库的分支系列,覆 盖最具知名度的 TOP50+ 门 AI 课程,旨在为读者和学习者提供一整套高品质中文 学习笔记和速查表。

点击课程名称,跳转至课程**资料包**页面,**一键下载**课程全部资料!

机器学习	深度学习	自然语言处理	计算机视觉
Stanford · CS229	Stanford · CS230	Stanford · CS224n	Stanford · CS23In

Awesome Al Courses Notes Cheatsheets· 持续更新中

知识图谱	图机器学习	深度强化学习	自动驾驶
Stanford · CS520	Stanford · CS224W	UCBerkeley · CS285	MIT · 6.S094

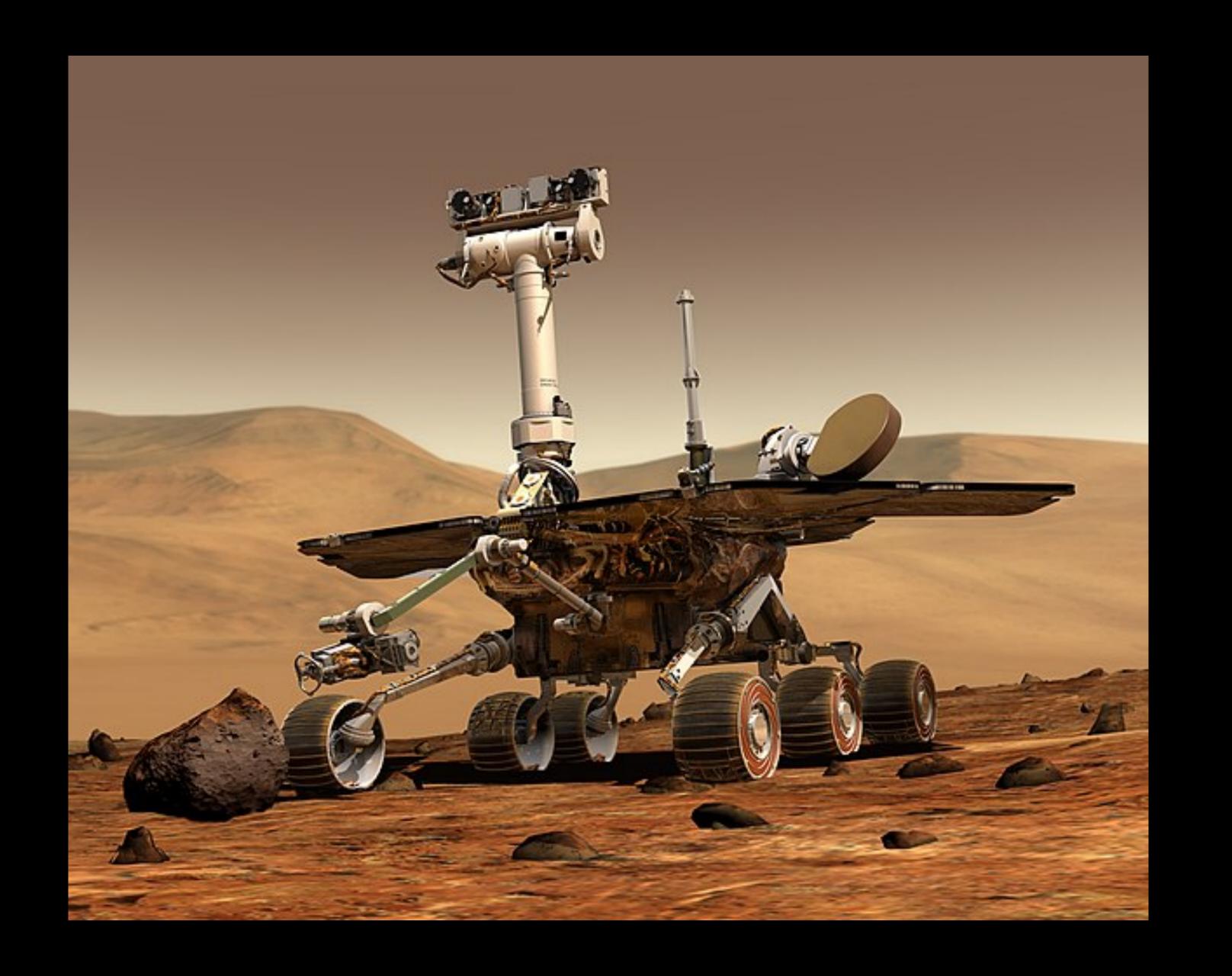


微信公众号

资料下载方式 2: 扫码点击底部菜单栏 称为 AI 内容创作者? 回复[添砖加瓦]

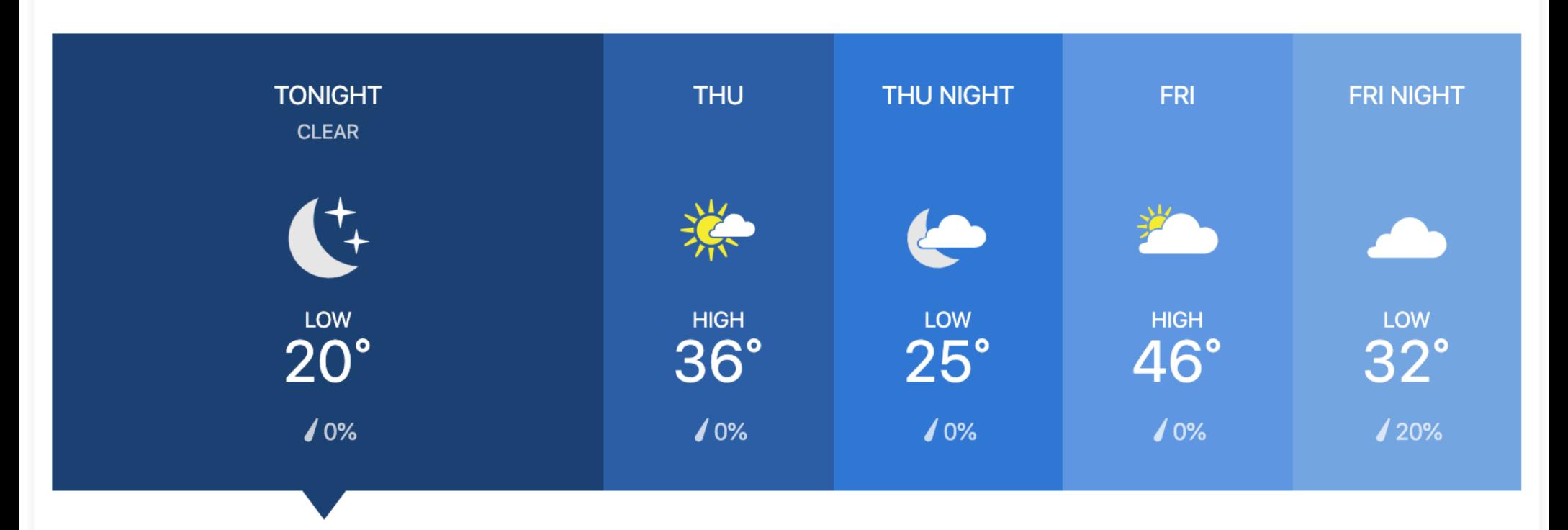
Artificial Intelligence with Python

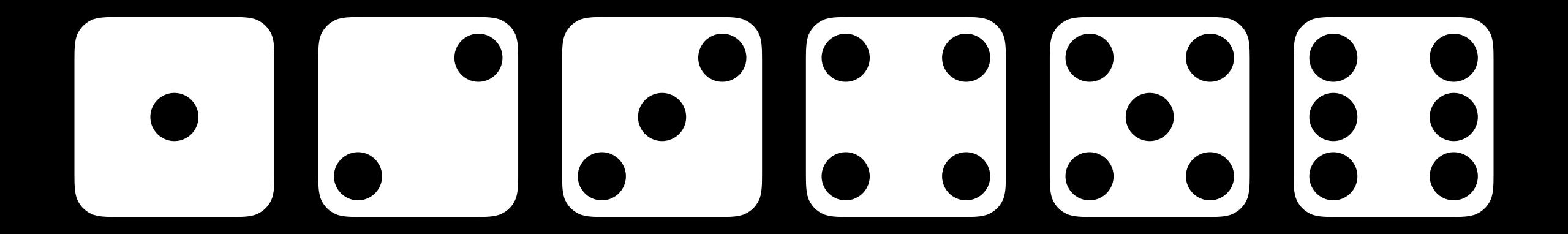
Uncertainty



NEXT 36 HOURS

HOURLY → 10 DAYS →





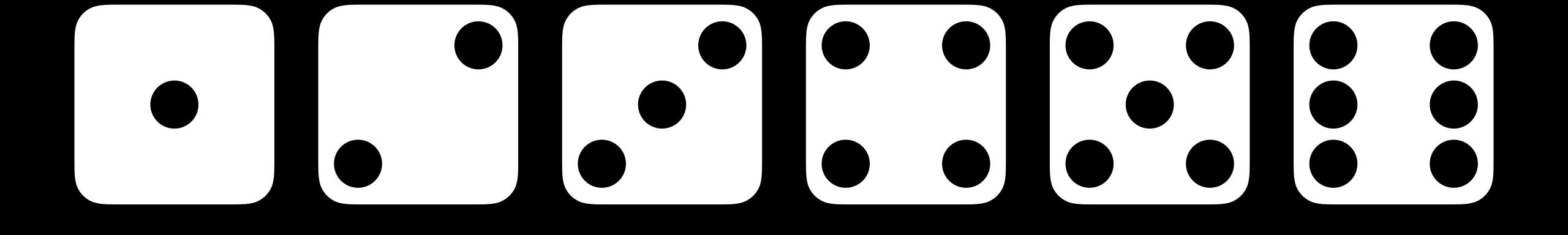
Probability

Possible Worlds



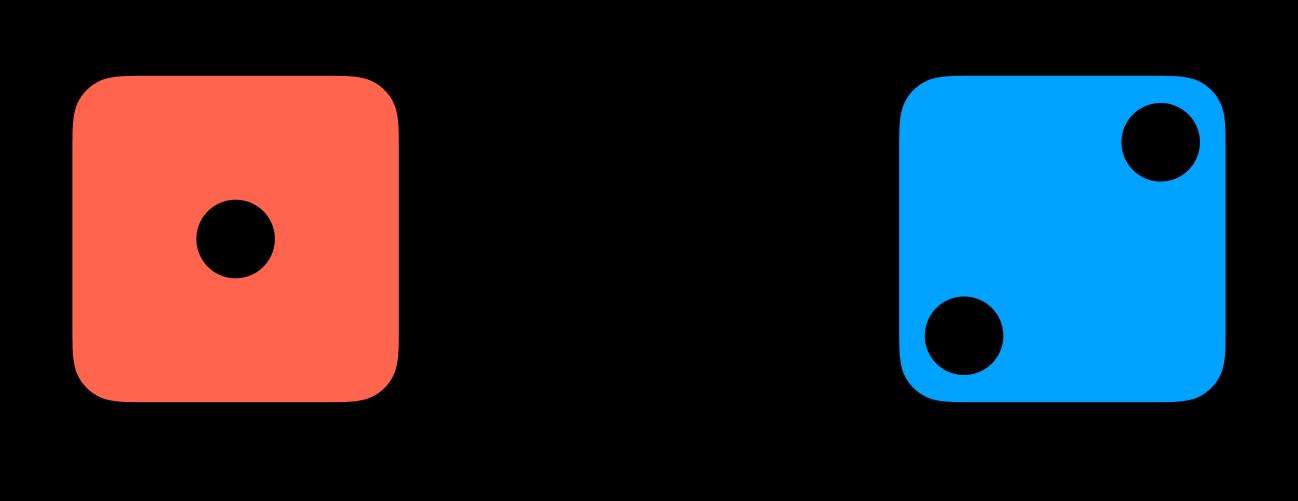
$0 \le P(\omega) \le 1$

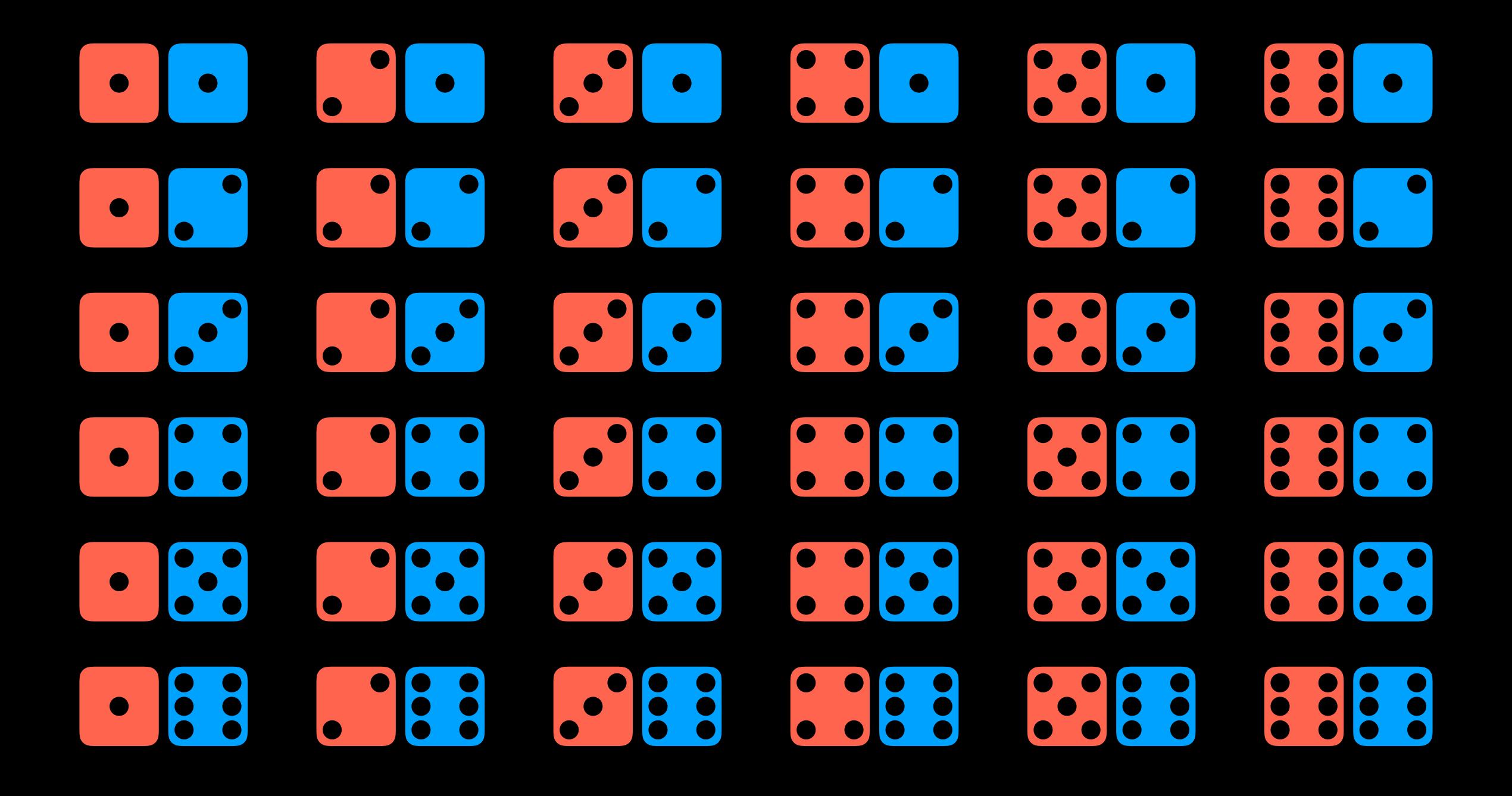
$\sum_{\omega \in \Omega} P(\omega) = 1$



6		6	6

$$P(6) = \frac{1}{6}$$











$P(sum\ to\ 12) = \frac{1}{36}$

$$P(sum\ to\ 7) = \frac{3}{36} = \frac{1}{6}$$

unconditional probability

degree of belief in a proposition in the absence of any other evidence

conditional probability

degree of belief in a proposition given some evidence that has already been revealed

conditional probability

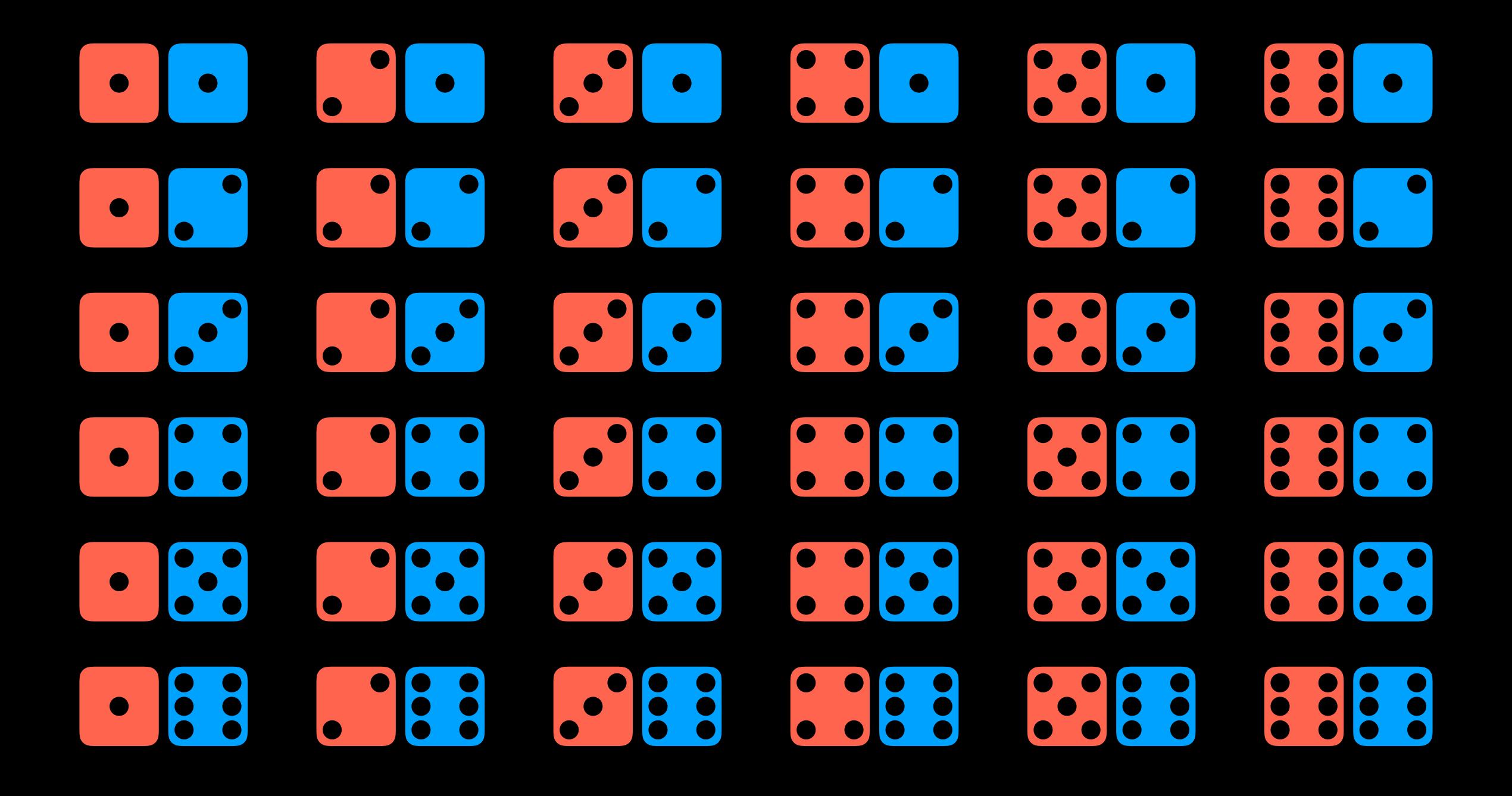
P(rain today | rain yesterday)

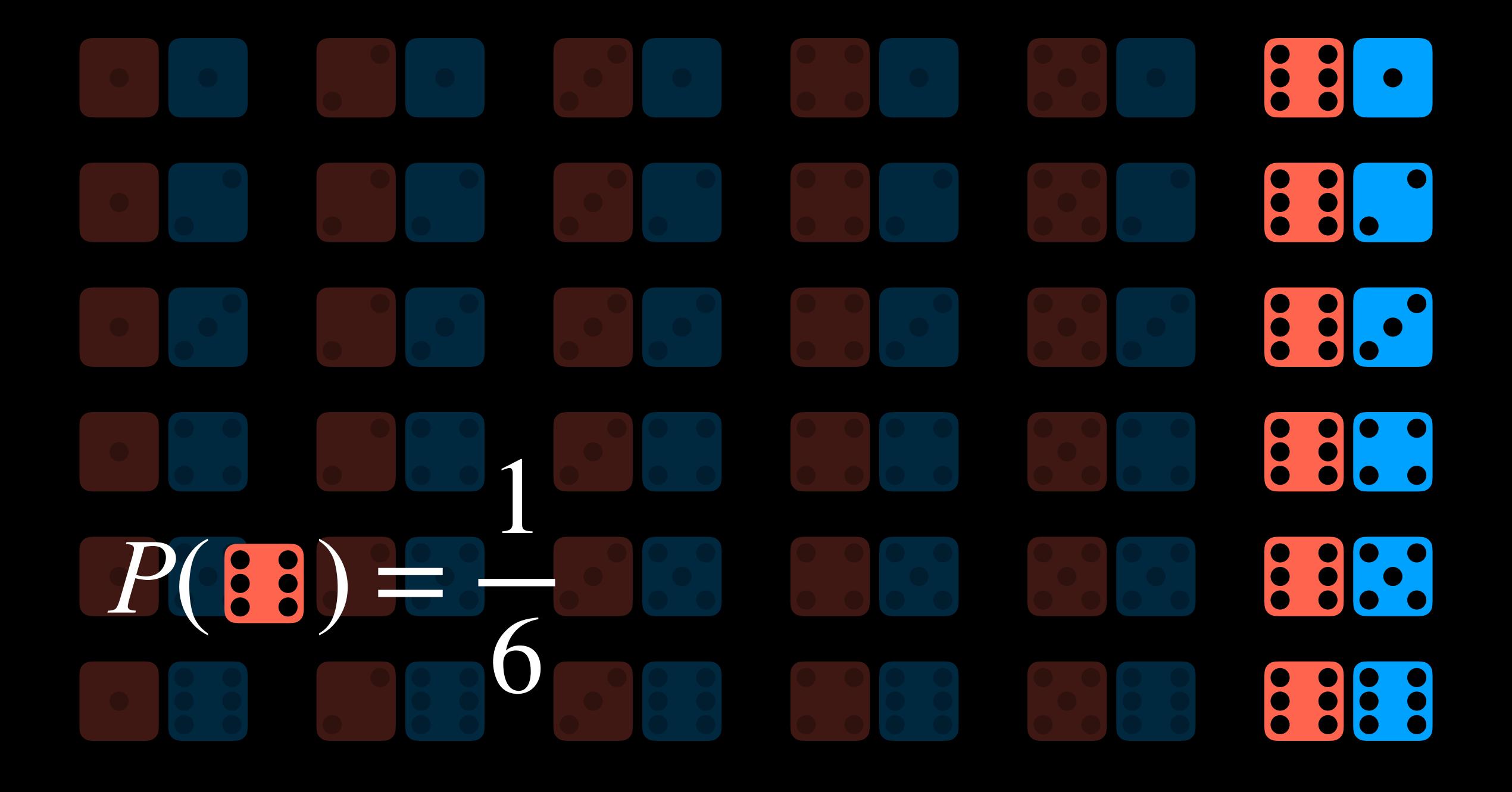
P(route change | traffic conditions)

P(disease | test results)

$P(a \mid b) = \frac{P(a \land b)}{P(b)}$

P(sum 12 | 88)





$$P(sum 12) = \frac{1}{36}$$
 $P(sum 12 | section 12 | section$

$$P(a \mid b) = \frac{P(a \land b)}{P(b)}$$

$$P(a \land b) = P(b)P(a \mid b)$$
$$P(a \land b) = P(a)P(b \mid a)$$

random variable

a variable in probability theory with a domain of possible values it can take on

random Variable

Roll

{1, 2, 3, 4, 5, 6}

random Variable

Weather

{sun, cloud, rain, wind, snow}

random variable

Traffic

{none, light, heavy}

random variable

Flight

{on time, delayed, cancelled}

probability distribution

$$P(Flight = on \ time) = 0.6$$

 $P(Flight = delayed) = 0.3$
 $P(Flight = cancelled) = 0.1$

probability distribution

$$P(Flight) = \langle 0.6, 0.3, 0.1 \rangle$$

the knowledge that one event occurs does not affect the probability of the other event

$$P(a \land b) = P(a)P(b \mid a)$$

$$P(a \land b) = P(a)P(b)$$

$$P(\blacksquare) = P(\blacksquare)P(\blacksquare)$$

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

$$P(\square) \neq P(\square)P(\square)$$

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

$$P(\square) \neq P(\square)P(\square)$$

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

Bayes' Rule

$$P(a \land b) = P(b) P(a \mid b)$$

$$P(a \land b) = P(a) P(b \mid a)$$

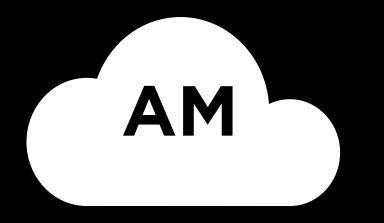
P(a) P(b | a) = P(b) P(a | b)

Bayes' Rule

$$P(b \mid a) = \frac{P(b) P(a \mid b)}{P(a)}$$

Bayes' Rule

$$P(b \mid a) = \frac{P(a \mid b) P(b)}{P(a)}$$





Given clouds in the morning, what's the probability of rain in the afternoon?

- 80% of rainy afternoons start with cloudy mornings.
- 40% of days have cloudy mornings.
- 10% of days have rainy afternoons.

$$P(rain | clouds) = \frac{P(clouds | rain)P(rain)}{P(clouds)}$$

$$= 0.2$$

P(cloudy morning | rainy afternoon)

we can calculate

P(rainy afternoon | cloudy morning)

P(visible effect | unknown cause)

we can calculate

P(unknown cause | visible effect)

P(medical test result | disease)

we can calculate

P(disease medical test result)

P(blurry text | counterfeit bill)

we can calculate

P(counterfeit bill | blurry text)

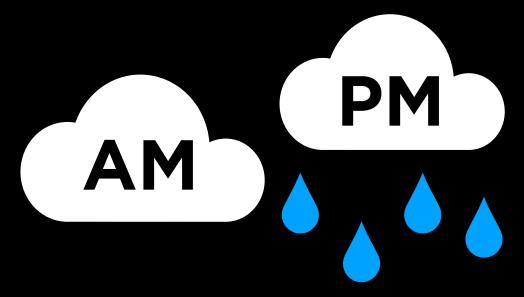
Joint Probability





C = cloud	$C = \neg cloud$
0.4	0.6

R = rain	$R = \neg rain$
0.1	0.9



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

$$\mathbf{P}(\mathbf{C} \mid \text{rain}) = \frac{\mathbf{P}(\mathbf{C}, \text{rain})}{\mathbf{P}(\text{rain})} = \alpha \mathbf{P}(\mathbf{C}, \text{rain})$$

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

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

Probability Rules

Negation

$$P(\neg a) = 1 - P(a)$$

Inclusion-Exclusion

$$P(a \lor b) = P(a) + P(b) - P(a \land b)$$

Marginalization

$$P(a) = P(a, b) + P(a, \neg b)$$

Marginalization

$$P(X = x_i) = \sum_{j} P(X = x_i, Y = y_j)$$

Marginalization

= 0.40

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

$$P(C = cloud)$$

$$= P(C = cloud, R = rain) + P(C = cloud, R = \neg rain)$$

$$= 0.08 + 0.32$$

Conditioning

$$P(a) = P(a | b)P(b) + P(a | \neg b)P(\neg b)$$

Conditioning

$$P(X = x_i) = \sum_{j} P(X = x_i | Y = y_j) P(Y = y_j)$$

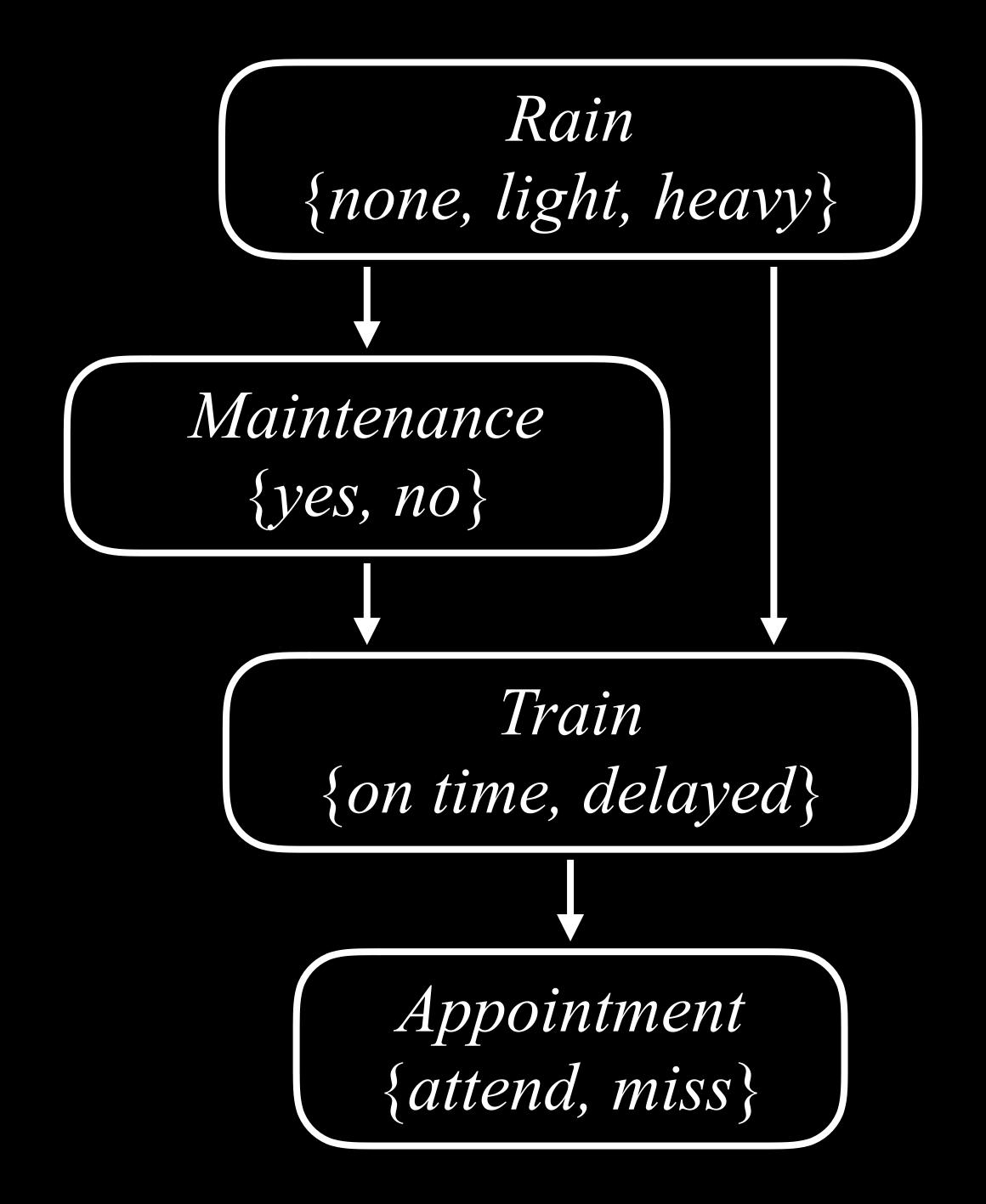
Bayesian Networks

Bayesian network

data structure that represents the dependencies among random variables

Bayesian network

- directed graph
- each node represents a random variable
- ullet arrow from X to Y means X is a parent of Y
- each node X has probability distribution $\mathbf{P}(X \mid Parents(X))$



Rain {none, light, heavy}

none	light	heavy
0.7	0.2	0.1

Rain {none, light, heavy}

Maintenance {yes, no}

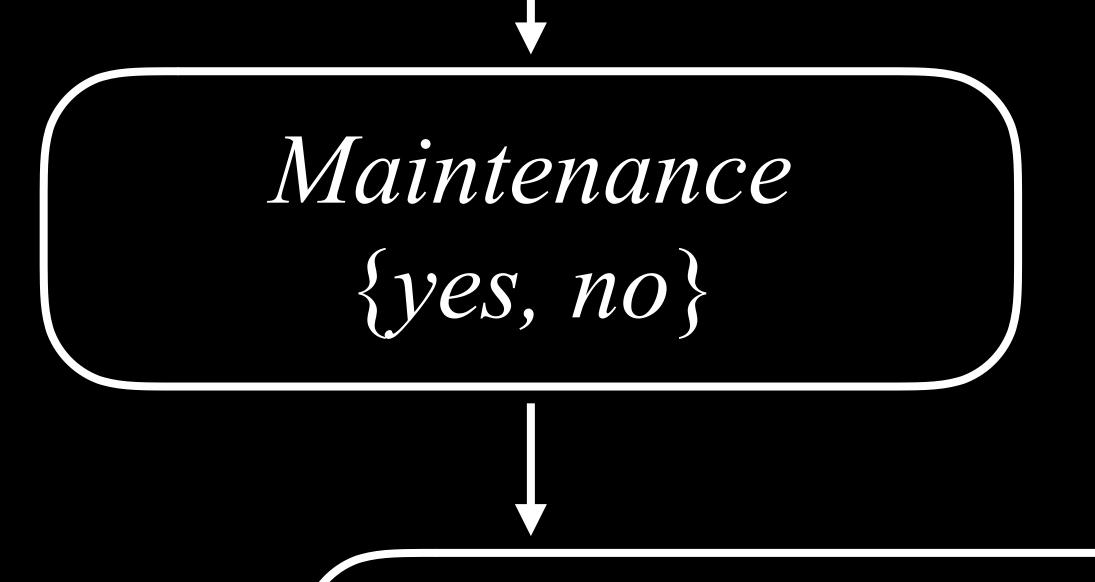
\boldsymbol{R}	yes	no
none	0.4	0.6
light	0.2	0.8
heavy	0.1	0.9

Rain {none, light, heavy}

Maintenance {yes, no}

Train
{on time, delayed}

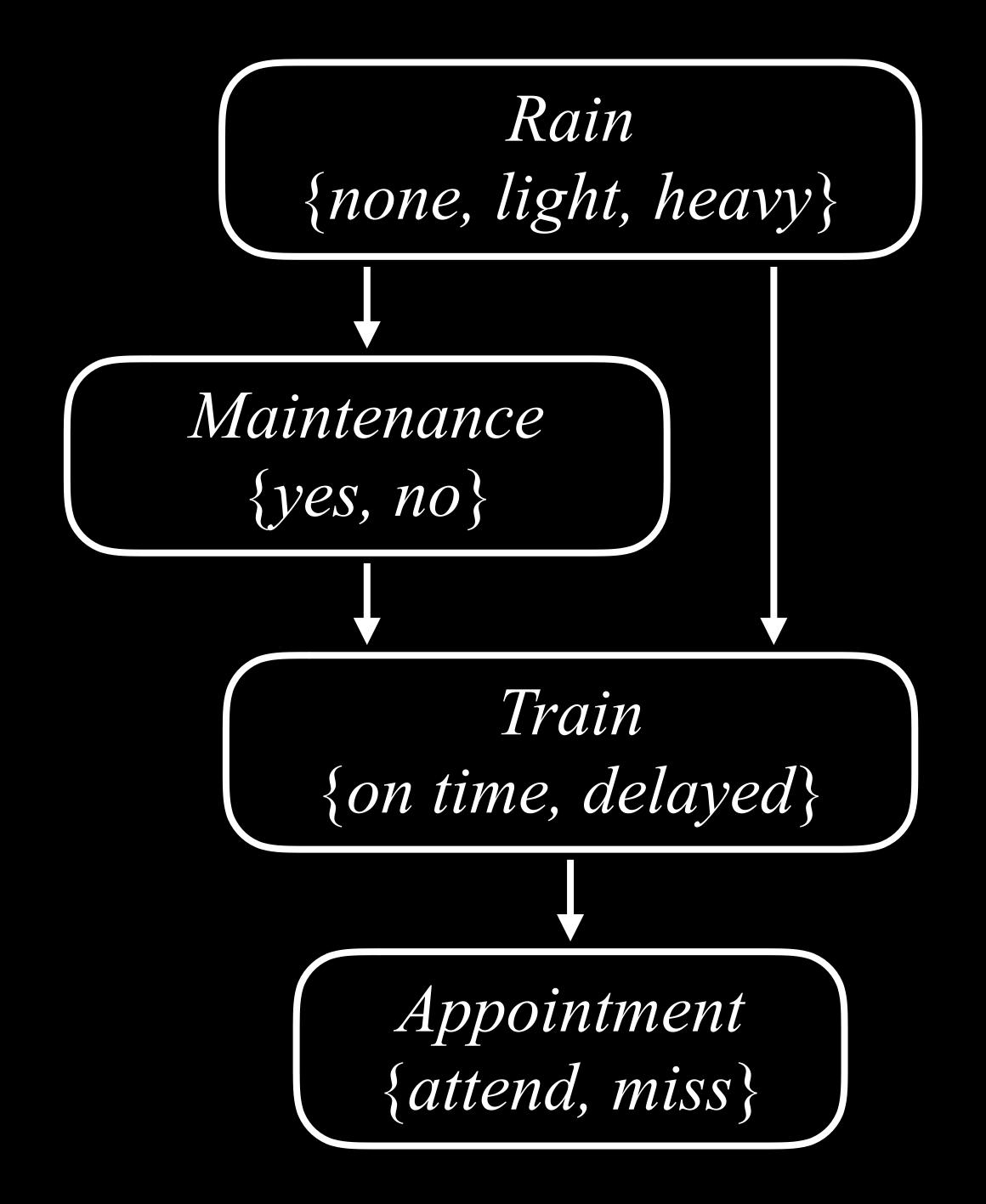
R	M	on time	delayed
none	yes	0.8	0.2
none	no	0.9	0.1
light	yes	0.6	0.4
light	no	0.7	0.3
heavy	yes	0.4	0.6
heavy	no	0.5	0.5

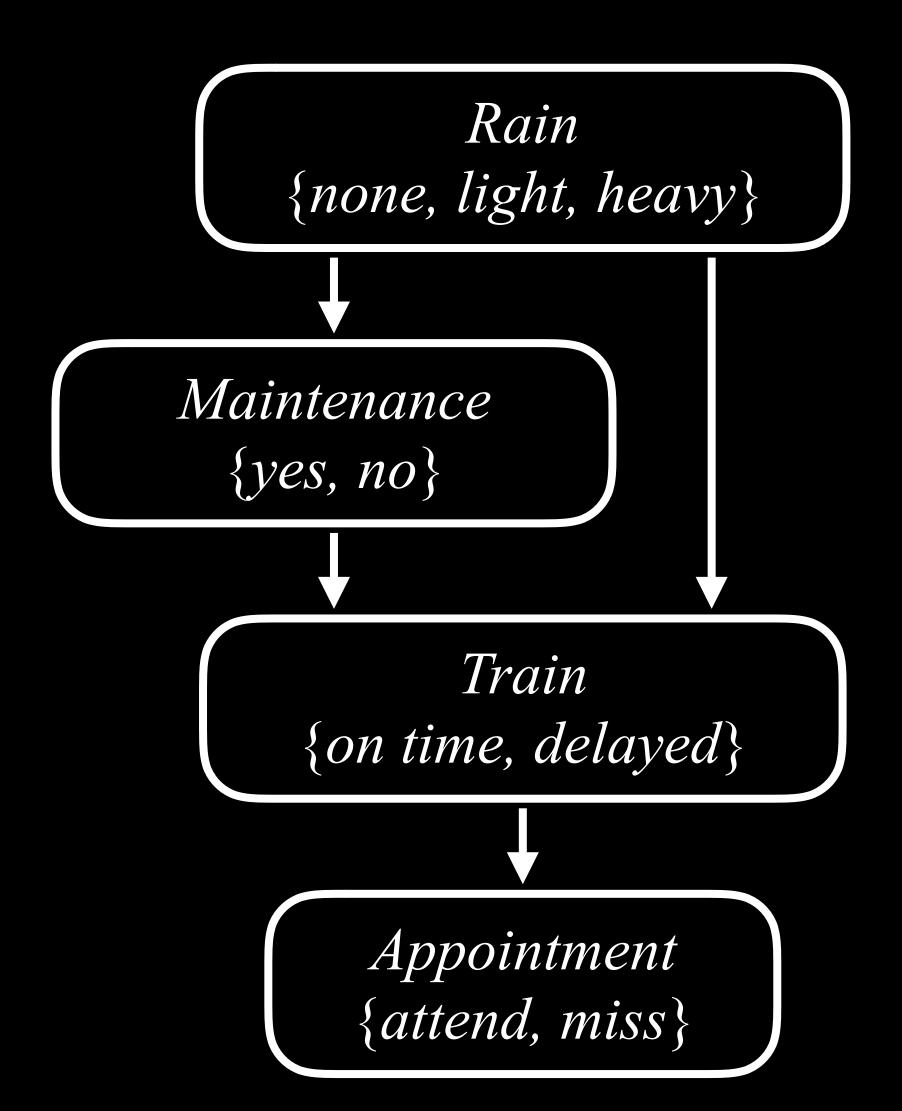


Train
{on time, delayed}

Appointment {attend, miss}

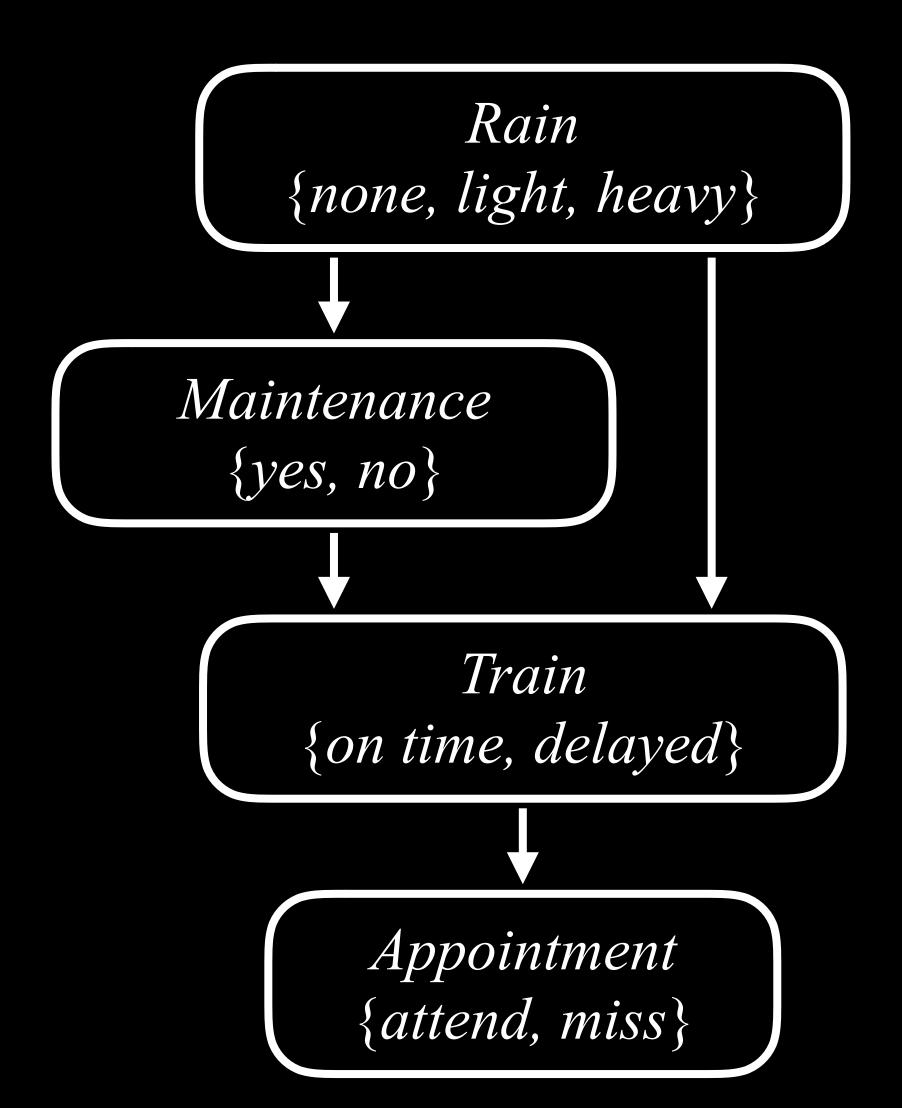
\boldsymbol{T}	attend	miss
on time	0.9	0.1
delayed	0.6	0.4





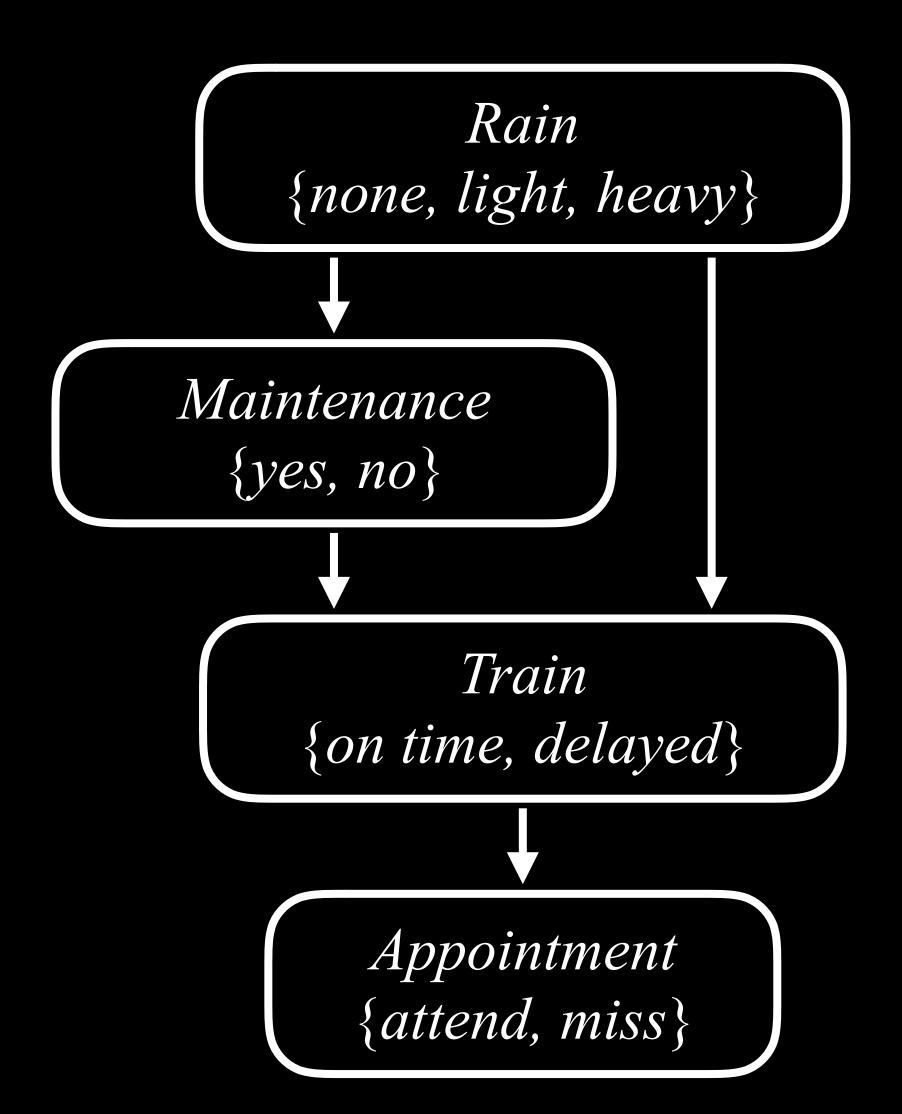
P(light)

P(light)



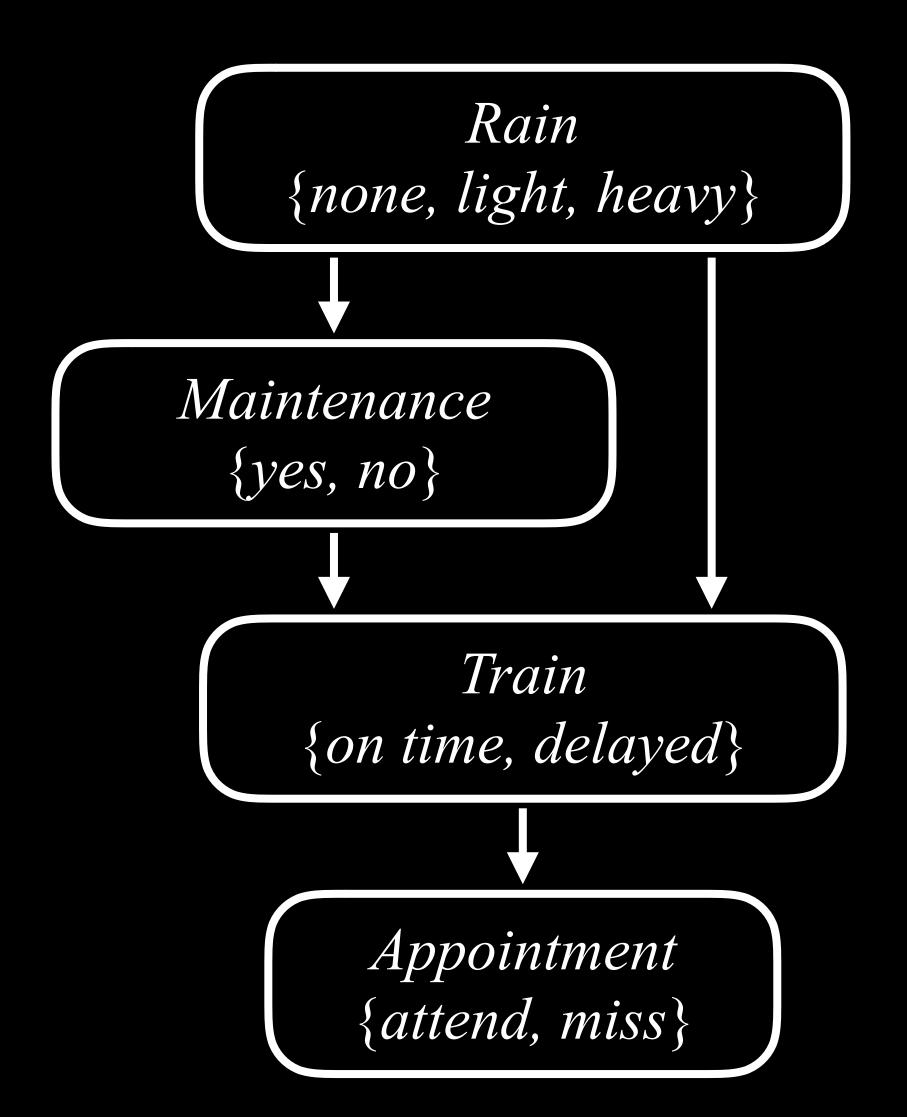
P(light, no)

P(light) P(no | light)



P(light, no, delayed)

 $P(light) P(no \mid light) P(delayed \mid light, no)$



P(light, no, delayed, miss)

 $P(light) P(no \mid light) P(delayed \mid light, no) P(miss \mid delayed)$

Inference

Inference

- Query X: variable for which to compute distribution
- Evidence variables E: observed variables for event e
- Hidden variables Y: non-evidence, non-query variable.

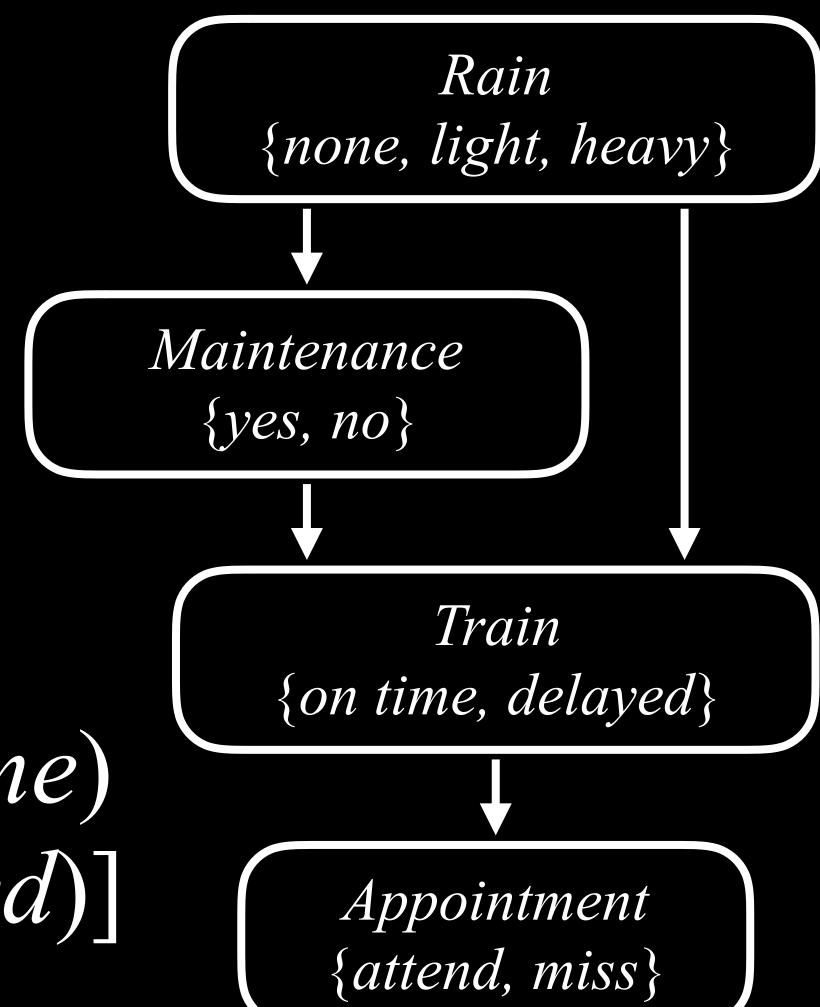
• Goal: Calculate P(X | e)

P(Appointment | light, no)

= α P(Appointment, light, no)

= α [P(Appointment, light, no, on time)

+ P(Appointment, light, no, delayed)]



Inference by Enumeration

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

X is the query variable.

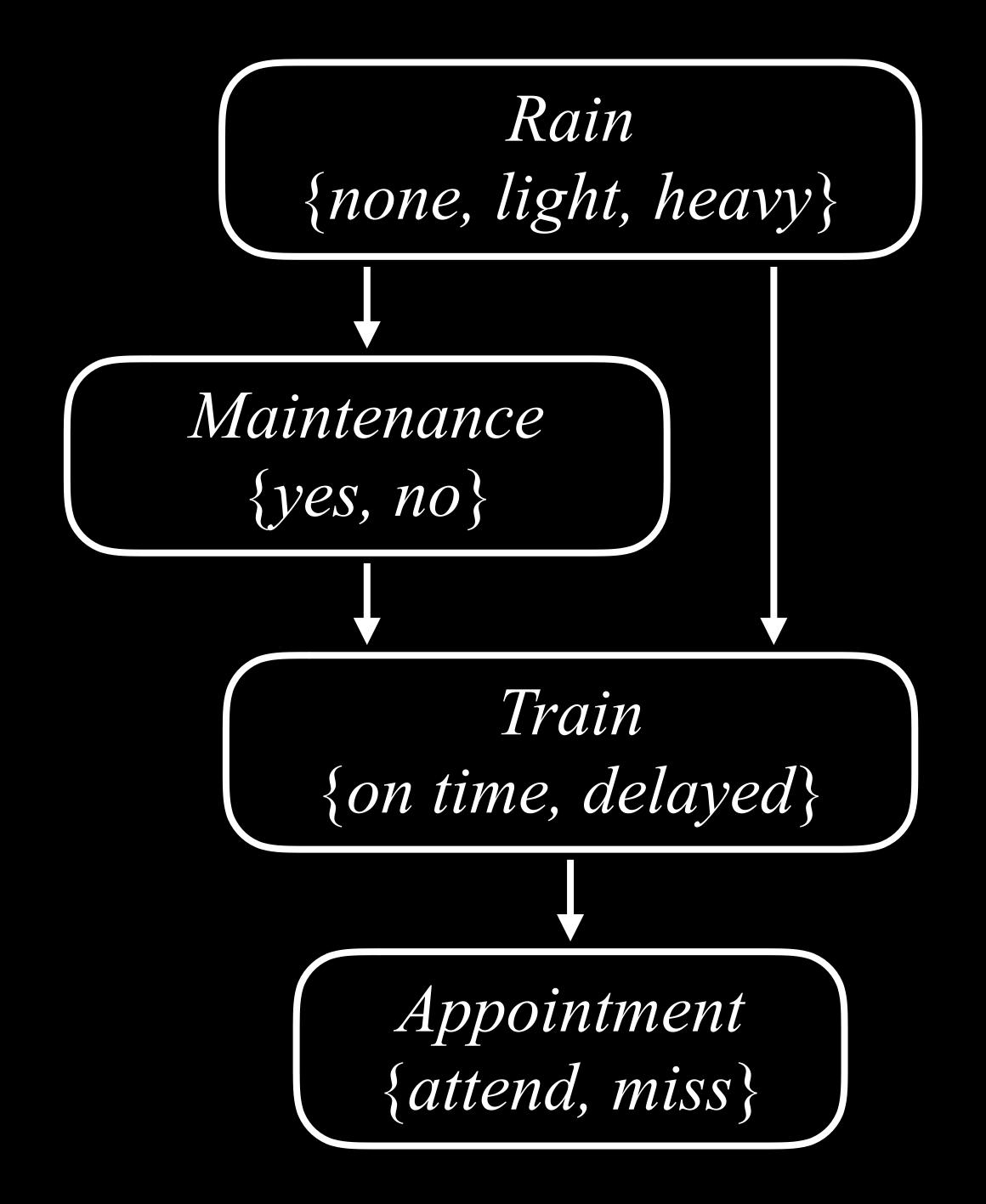
e is the evidence.

y ranges over values of hidden variables.

α normalizes the result.

Approximate Inference

Sampling



R = none	

Rain
{none, light, heavy}

none	light	heavy
0.7	0.2	0.1

$$R = none$$
 $M = yes$

Rain {none, light, heavy}

Maintenance {ves, no}

\boldsymbol{R}	yes	no
none	0.4	0.6
light	0.2	0.8
heavy	0.1	0.9

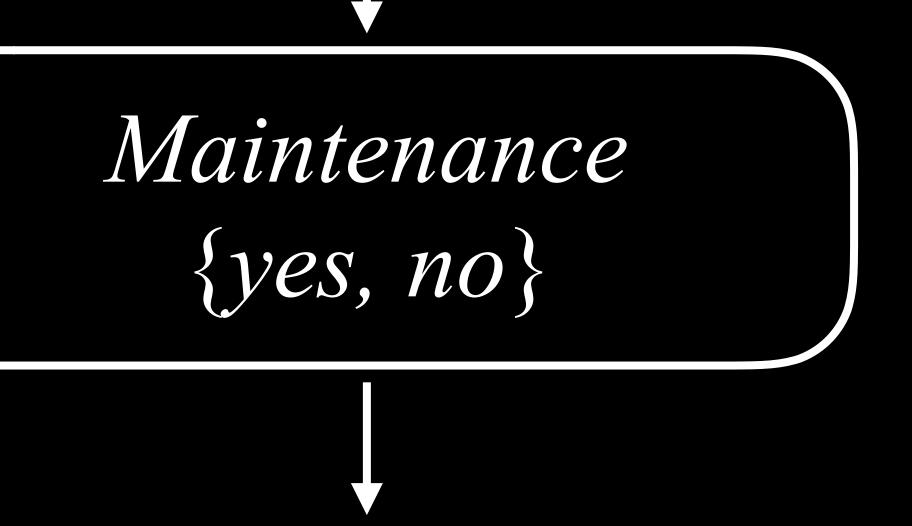
Rain {none, light, heavy}

Maintenance {yes, no}

Train
{on time, delayed}

R = none M = yes T = on time

R	M	on time	delayed
none	yes	0.8	0.2
none	no	0.9	0.1
light	yes	0.6	0.4
light	no	0.7	0.3
heavy	yes	0.4	0.6
heavy	no	0.5	0.5



Train
{on time, delayed}

R = none
M = yes
T = on time
A = attend

Appointment {attend, miss}

\boldsymbol{T}	attend	miss
on time	0.9	0.1
delayed	0.6	0.4

R = none

M = yes

T = on time

A = attend

R = li	ght
M =	no
$\Gamma = on$	time

$$R = light$$

$$M = yes$$

$$T = delayed$$

$$A = attend$$

$$R = none$$

$$M = no$$

$$T = on time$$

$$A = attend$$

$$R = none$$

$$M = yes$$

$$T = on time$$

$$A = attend$$

$$R = none$$

A = miss

$$M = yes$$

$$T = on time$$

$$A = attend$$

$$R = none$$

$$M = yes$$

$$T = on time$$

$$A = attend$$

$$R = heavy$$

$$M = no$$

$$T = delayed$$

$$A = miss$$

$$R = light$$

$$M = no$$

$$T = on time$$

$$A = attend$$

P(Train = on time)?

R = li	ght
M =	no
$\Gamma = on$	time

$$R = light$$

$$M = yes$$

$$T = delayed$$

$$A = attend$$

$$R = none$$

$$M = no$$

$$T = on time$$

$$A = attend$$

$$R = none$$

$$M = yes$$

$$T = on time$$

$$A = attend$$

$$R = none$$

A = miss

$$M = yes$$

$$T = on time$$

$$A = attend$$

$$R = none$$

$$M = yes$$

$$T = on time$$

$$A = attend$$

$$R = heavy$$

$$M = no$$

$$T = delayed$$

$$A = miss$$

$$R = light$$

$$M = no$$

$$T = on time$$

$$A = attend$$

R = light
M = no
T = on time
A = miss

$$R = light$$

$$M = yes$$

$$T = delayed$$

$$A = attend$$

$$R = none$$

$$M = no$$

$$T = on time$$

$$A = attend$$

$$R = none$$

$$M = yes$$

$$T = on time$$

$$A = attend$$

$$R = none$$

$$M = yes$$

$$T = on time$$

$$A = attend$$

$$R = none$$

$$M = yes$$

$$T = on time$$

$$A = attend$$

$$R = heavy$$

$$M = no$$

$$T = delayed$$

$$A = miss$$

$$R = light$$

$$M = no$$

$$T = on time$$

$$A = attend$$

P(Rain = light | Train = on time)?

R = light
M = no
$\Gamma = on time$

$$R = light$$

$$M = yes$$

$$T = delayed$$

$$A = attend$$

$$R = none$$

$$M = no$$

$$T = on time$$

$$A = attend$$

$$R = none$$

$$M = yes$$

$$T = on time$$

$$A = attend$$

$$R = none$$

A = miss

$$M = yes$$

$$T = on time$$

$$A = attend$$

$$R = none$$

$$M = yes$$

$$T = on time$$

$$A = attend$$

$$R = heavy$$

$$M = no$$

$$T = delayed$$

$$A = miss$$

$$R = light$$

$$M = no$$

$$T = on time$$

$$A = attend$$

D	1. 1
K =	llght
	118111

M = no

T = on time

A = miss

R = light

M = yes

T = delayed

A = attend

R = none

M = no

T = on time

A = attend

R = none

M = yes

T = on time

A = attend

R = none

M = yes

T = on time

A = attend

R = none

M = yes

T = on time

A = attend

R = heavy

M = no

T = delayed

A = miss

R = light

M = no

T = on time

A = attend

$$R = light$$

M = no

T = on time

A = miss

R = light

M = yes

T = delayed

A = attend

R = none

M = no

T = on time

A = attend

R = none

M = yes

T = on time

A = attend

R = none

M = yes

T = on time

A = attend

R = none

M = yes

T = on time

A = attend

R = heavy

M = no

T = delayed

A = miss

R = light

M = no

T = on time

A = attend

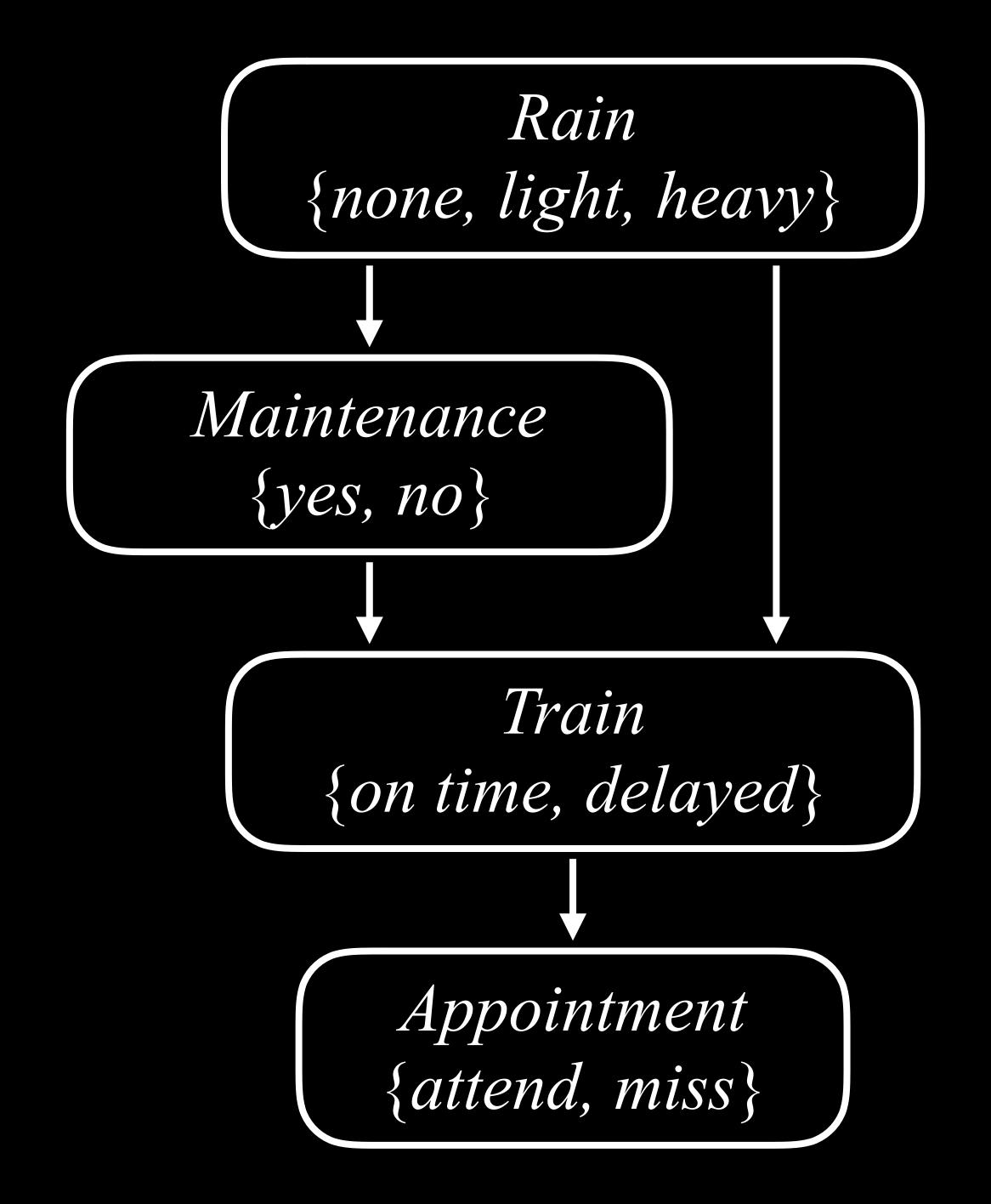
Rejection Sampling

Likelihood Weighting

Likelihood Weighting

- Start by fixing the values for evidence variables.
- Sample the non-evidence variables using conditional probabilities in the Bayesian Network.
- Weight each sample by its **likelihood**: the probability of all of the evidence.

P(Rain = light | Train = on time)?



R = light

T = on time

Rain {none, light, heavy}

none	light	heavy
0.7	0.2	0.1

$$R = light$$

$$M = yes$$

$$T = on time$$

Rain {none, light, heavy}

Maintenance {yes, no}

\boldsymbol{R}	yes	no
none	0.4	0.6
light	0.2	0.8
heavy	0.1	0.9

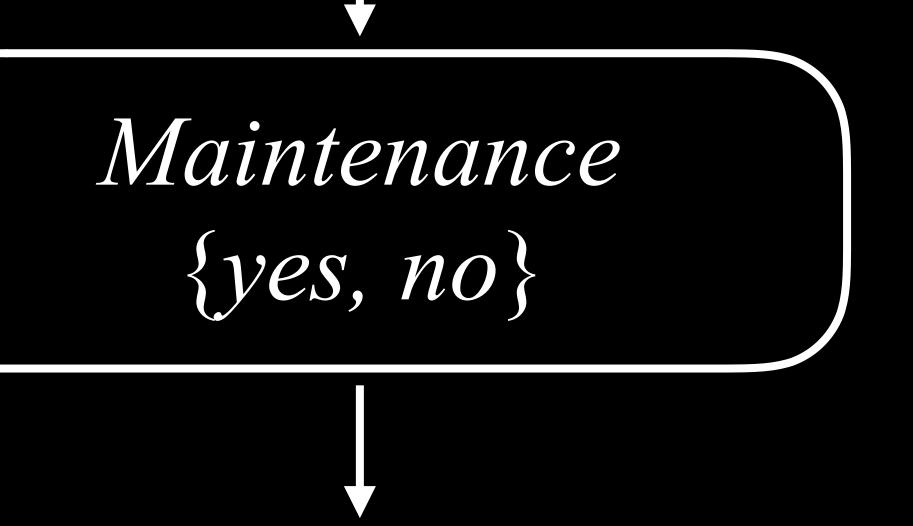
Rain {none, light, heavy}

R = light M = yes T = on time

Maintenance {yes, no}

Train
{on time, delayed}

R	M	on time	delayed
none	yes	0.8	0.2
none	no	0.9	0.1
light	yes	0.6	0.4
light	no	0.7	0.3
heavy	yes	0.4	0.6
heavy	no	0.5	0.5



Train
{on time, delayed}

R = light
M = yes
T = on time
A = attend

Appointment {attend, miss}

T	attend	miss
on time	0.9	0.1
delayed	0.6	0.4

Rain {none, light, heavy}

Maintenance {yes, no}

Train
{on time, delayed}

R = light
M = yes
T = on time
A = attend

R	M	on time	delayed
none	yes	0.8	0.2
none	no	0.9	0.1
light	yes	0.6	0.4
light	no	0.7	0.3
heavy	yes	0.4	0.6
heavy	no	0.5	0.5

Rain {none, light, heavy}

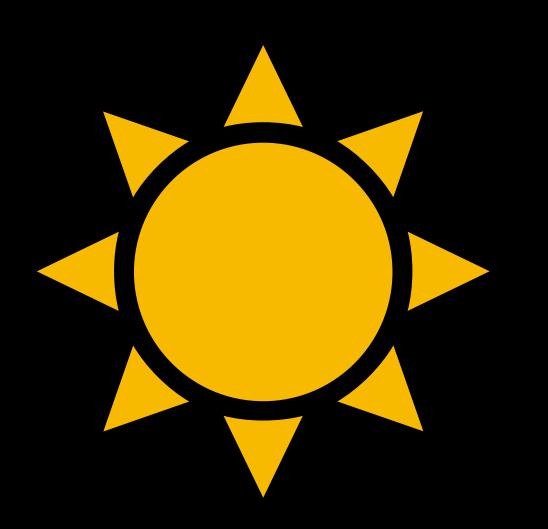
Maintenance {yes, no}

Train
{on time, delayed}

R = light M = yes T = on time A = attend

R	M	on time	delayed
none	yes	0.8	0.2
none	no	0.9	0.1
light	yes	0.6	0.4
light	no	0.7	0.3
heavy	yes	0.4	0.6
heavy	no	0.5	0.5

Uncertainty over Time





Xt: Weather at time t

Markov assumption

the assumption that the current state depends on only a finite fixed number of previous states

Markov Chain

Markov chain

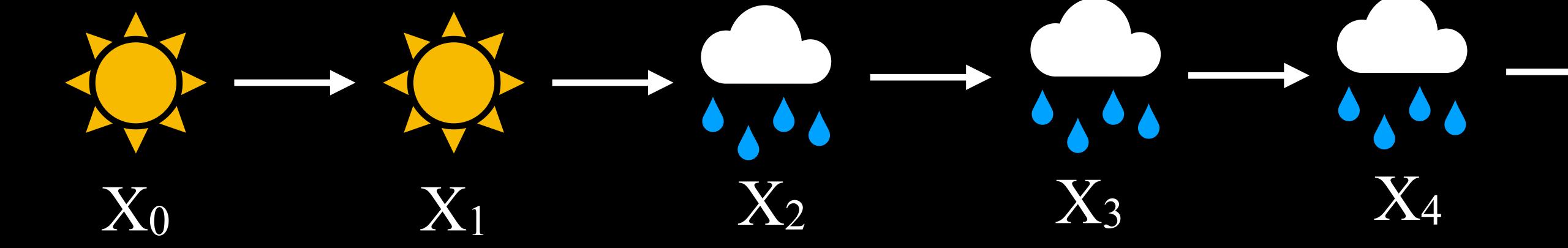
a sequence of random variables where the distribution of each variable follows the Markov assumption

Transition Mode

Tomorrow (X_{t+1})

0.8	0.2
0.3	0.7

Today (X_t)



Sensor Models

Hidden State	Observation	
robot's position	robot's sensor data	
words spoken	audio waveforms	
user engagement	website or app analytics	
weather	umbrella	

Hidden Markov Models

Hidden Markov Model

a Markov model for a system with hidden states that generate some observed event

Sensor Mode

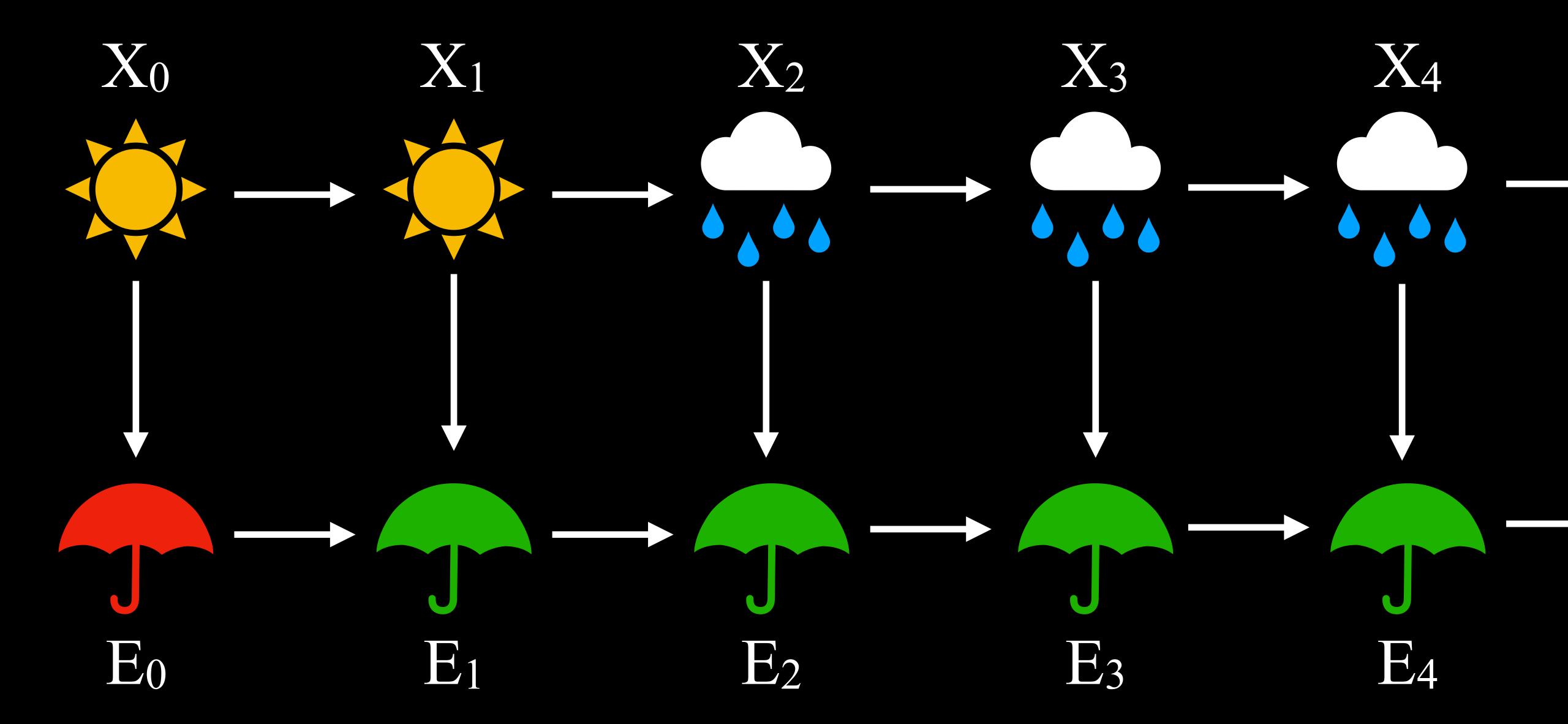
Observation (E_t)

0.2	0.8
0.9	0.1

State (X_t)

sensor Markov assumption

the assumption that the evidence variable depends only the corresponding state



Task	Definition
filtering	given observations from start until now, calculate distribution for current state
prediction	given observations from start until now, calculate distribution for a future state
smoothing	given observations from start until now, calculate distribution for past state
most likely explanation	given observations from start until now, calculate most likely sequence of states

Uncertainty

Artificial Intelligence with Python

Harvard · CS50-Al | Introduction to Artificial Intelligence with Python (2020)

CS50-AI (2020)· 课程资料包 @ShowMeAI









视频 中英双语字幕

课件 一键打包下载

筆记 官方笔记翻译

代码 作业项目解析



视频·B站[扫码或点击链接]

https://www.bilibili.com/video/BV1AQ4y1y7wy



课件 & 代码·博客[扫码或点击链接]

http://blog.showmeai.tech/harvard-cs50-ai

RNN

深度优先 广度优先

A* 算法 贝叶斯 自然语言处理

聚类 马尔可夫

强化学习 tensorflow 反向传播

SVM

监督学习

无监督

语法解析

神经网络 CNN

搜索

Awesome Al Courses Notes Cheatsheets 是 **ShowMeAI** 资料库的分支系列,覆 盖最具知名度的 TOP50+ 门 AI 课程,旨在为读者和学习者提供一整套高品质中文 学习笔记和速查表。

点击课程名称,跳转至课程**资料包**页面,**一键下载**课程全部资料!

机器学习	深度学习	自然语言处理	计算机视觉
Stanford · CS229	Stanford · CS230	Stanford · CS224n	Stanford · CS23In

Awesome Al Courses Notes Cheatsheets· 持续更新中

知识图谱	图机器学习	深度强化学习	自动驾驶
Stanford · CS520	Stanford · CS224W	UCBerkeley · CS285	MIT · 6.S094



微信公众号

资料下载方式 2: 扫码点击底部菜单栏 称为 AI 内容创作者? 回复[添砖加瓦]