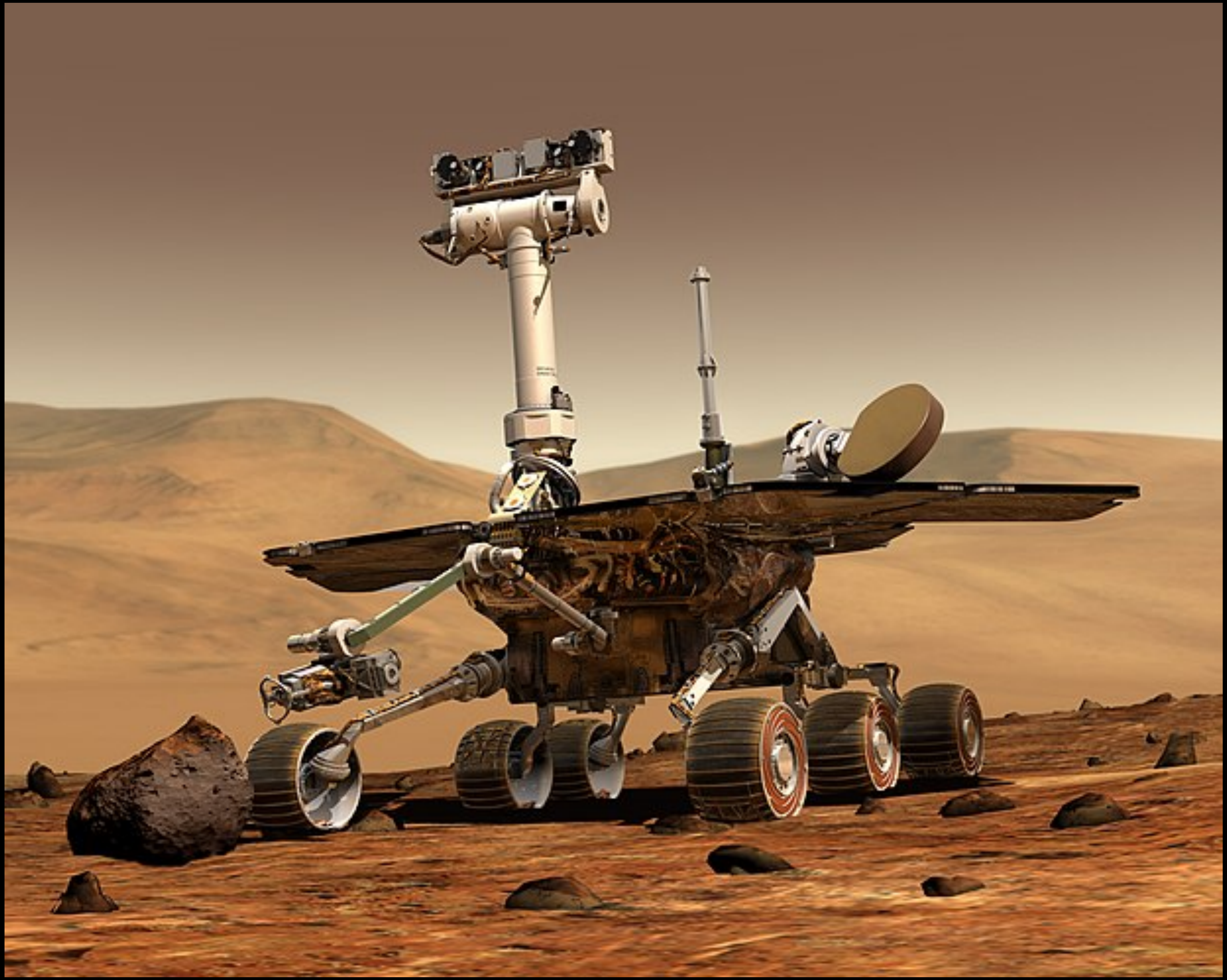
Introduction to

Artificial Intelligence

with Python

YouTube Link:
https://youtu.be/D8RRq3TbtHU?list=PLhQjrBD2T381PopUTYtMSstgk-hsTGkVm

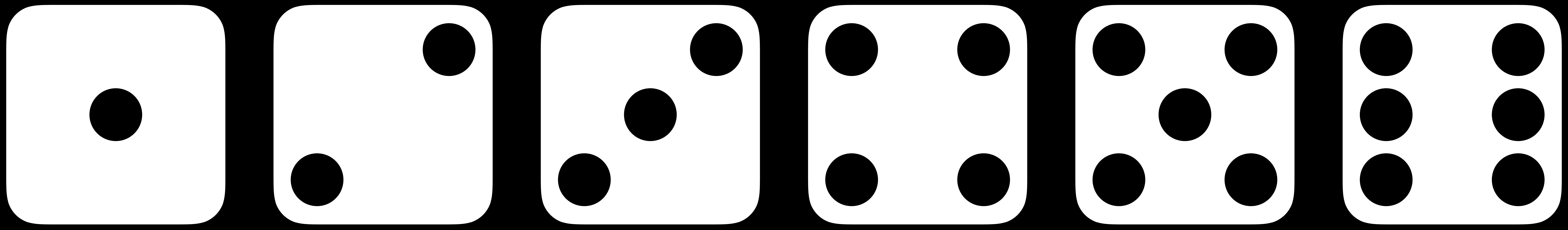
Uncertainty



NEXT 36 HOURS

HOURLY → | 10 DAYS →

TONIGHT	THU	THU NIGHT	FRI	FRI NIGHT
CLEAR				
LOW	HIGH	LOW	HIGH	LOW
20°	36°	25°	46°	32°
0%	0%	0%	0%	20%



Probability

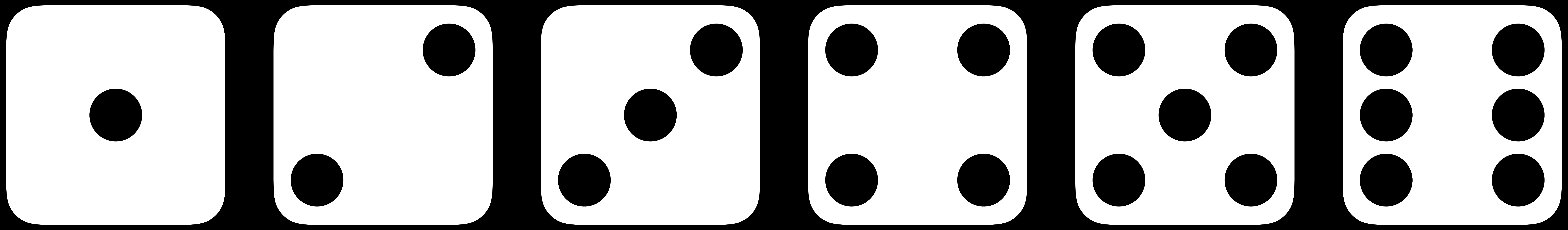
Possible Worlds

$$\omega$$

$$P(\omega)$$

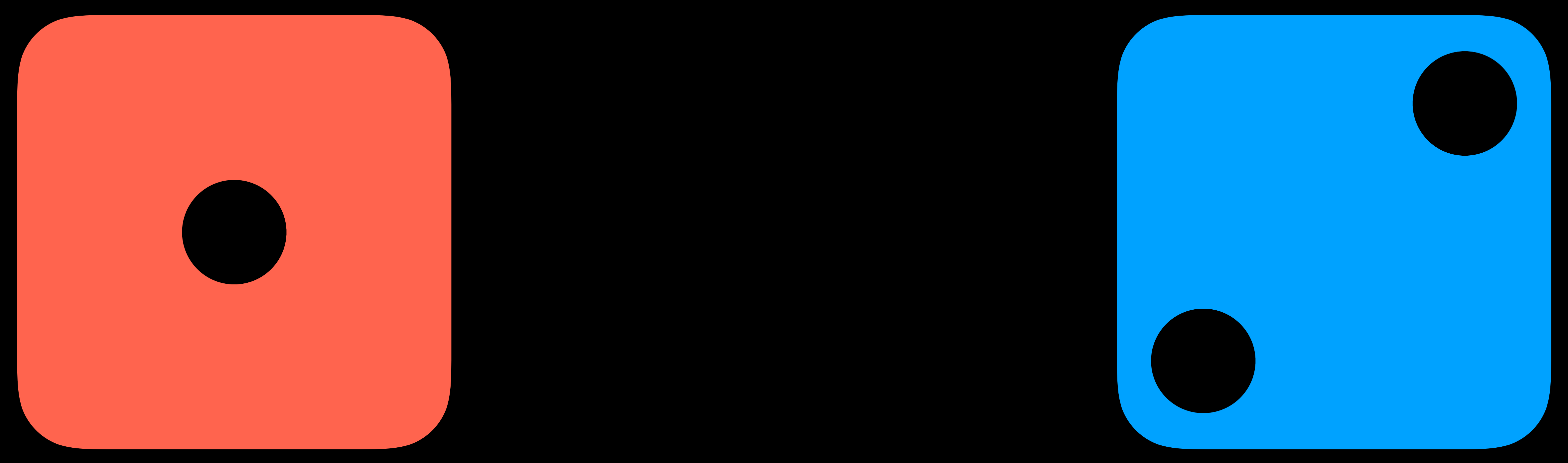
$$0 \leq P(\omega) \leq 1$$

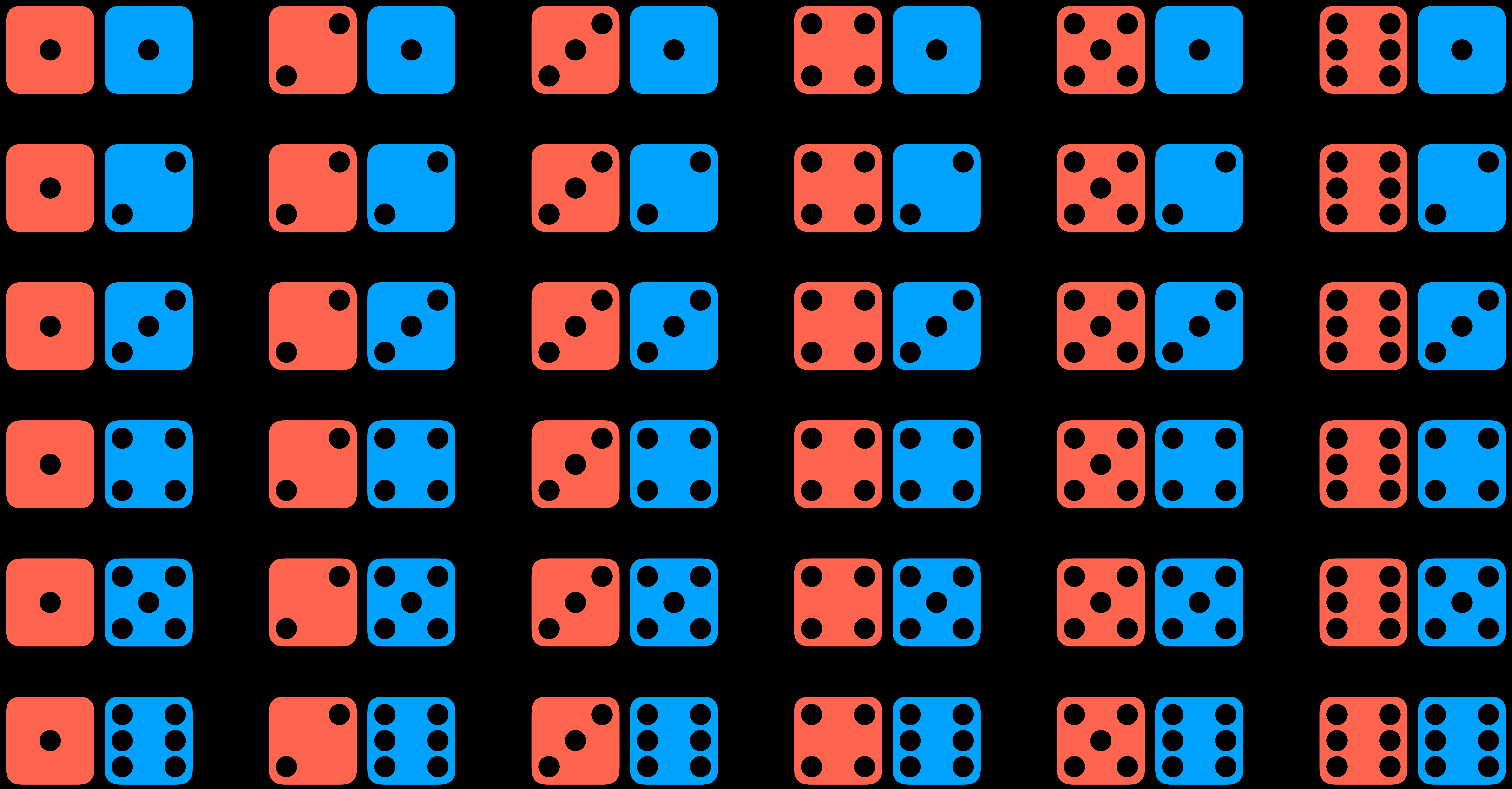
$$\sum_{\omega \in \Omega} P(\omega) = 1$$



$\frac{1}{6}$ $\frac{1}{6}$ $\frac{1}{6}$ $\frac{1}{6}$ $\frac{1}{6}$ $\frac{1}{6}$

$$P(\text{⚁}) = \frac{1}{6}$$





2	3	4	5	6	7
3	4	5	6	7	8
4	5	6	7	8	9
5	6	7	8	9	10
6	7	8	9	10	11
7	8	9	10	11	12

2	3	4	5	6	7
3	4	5	6	7	8
4	5	6	7	8	9
5	6	7	8	9	10
6	7	8	9	10	11
7	8	9	10	11	12

2	3	4	5	6	7
3	4	5	6	7	8
4	5	6	7	8	9
5	6	7	8	9	10
6	7	8	9	10	11
7	8	9	10	11	12

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

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

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

$$P(a \mid b)$$

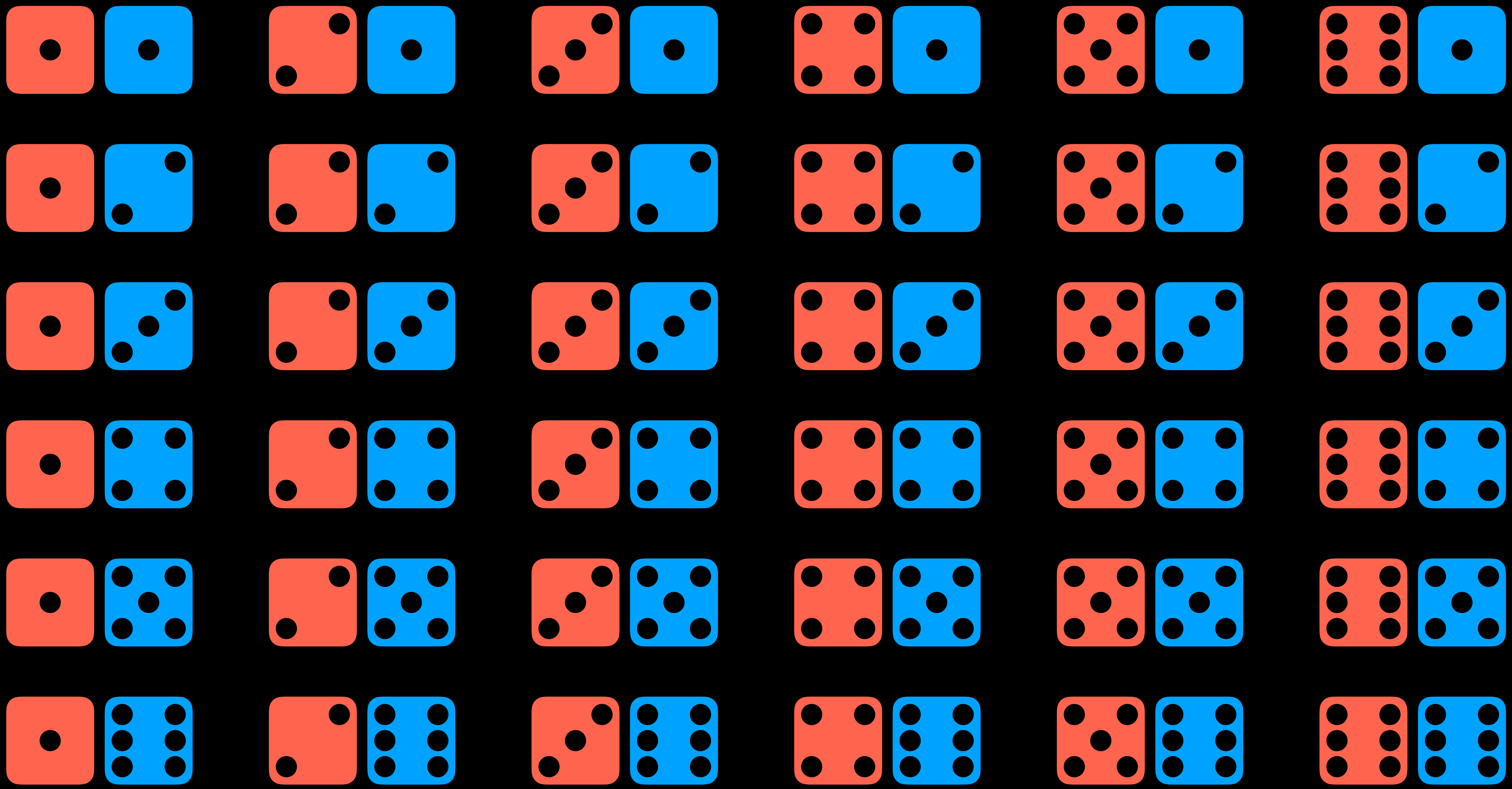
$$P(rain\ today \mid rain\ yesterday)$$

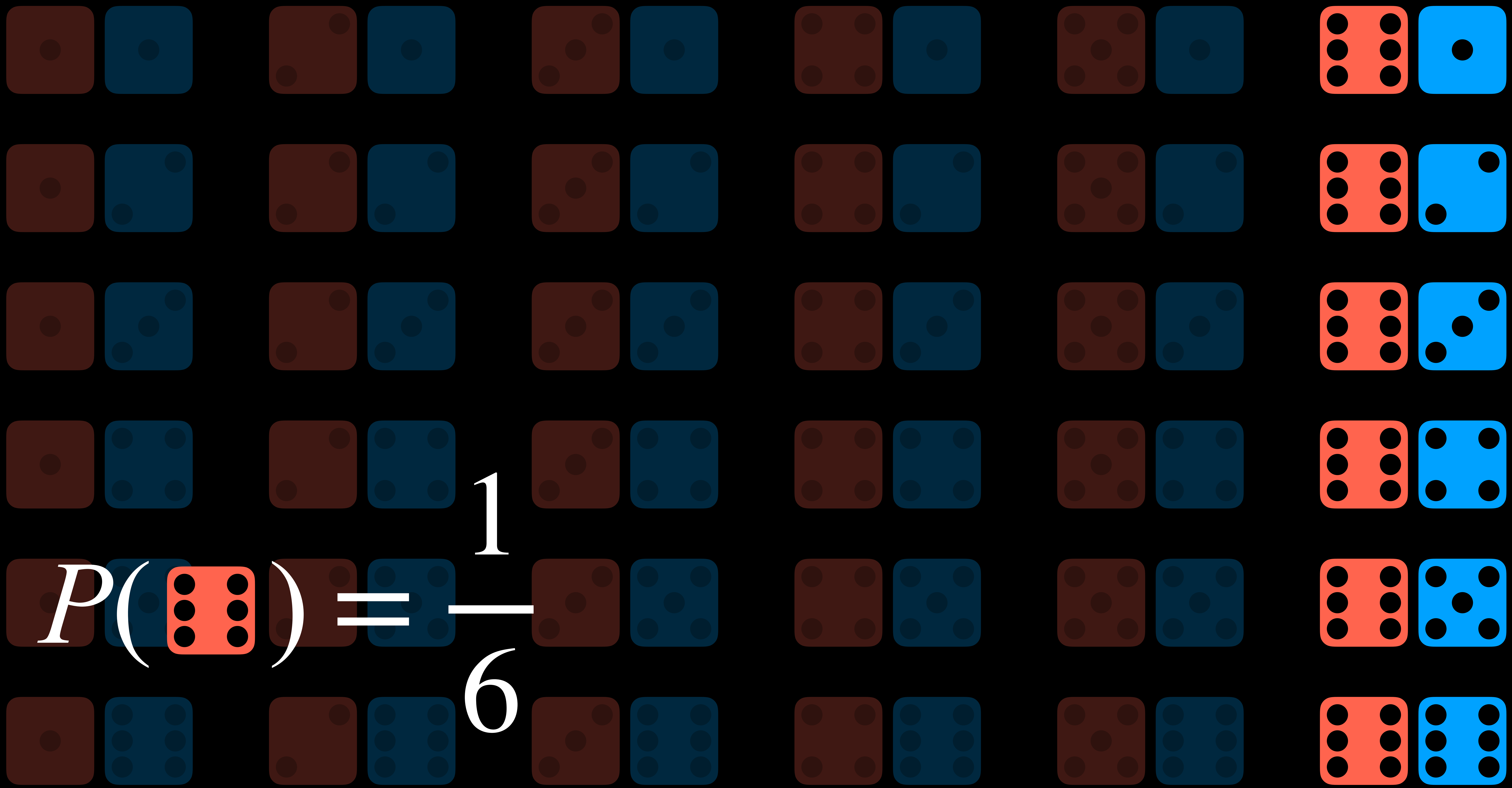
$$P(\mathit{route\ change} \mid \mathit{traffic\ conditions})$$

$$P(disease \mid test\ results)$$

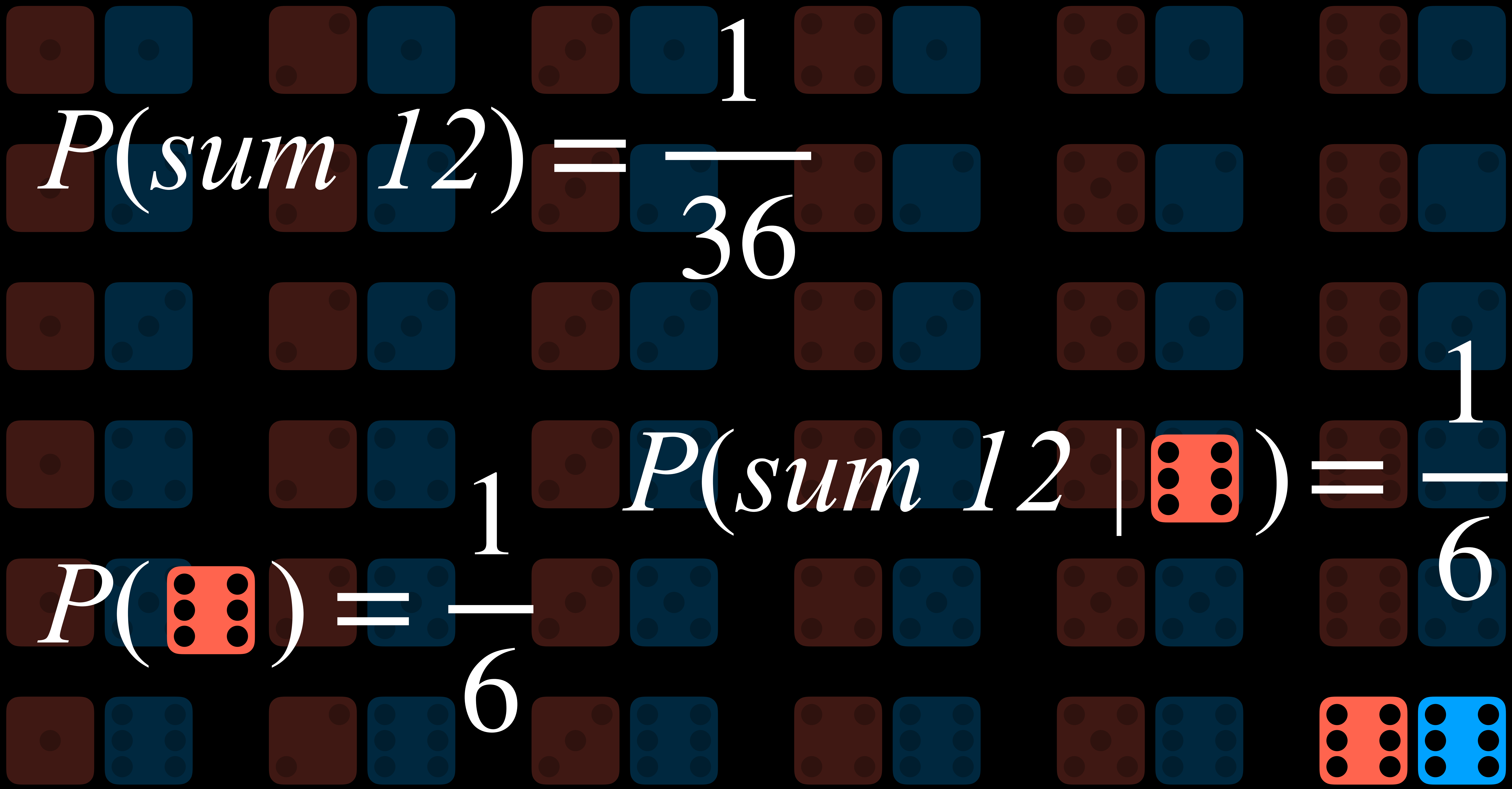
$$P(a \mid b) = \frac{P(a \wedge b)}{P(b)}$$

$P(sum\ 12 \mid$ ⚅ $)$





$$P(\text{⚅}) = \frac{1}{6}$$



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

$$P(sum\ 12 \mid ⚅) = \frac{1}{6}$$

$$P(⚅) = \frac{1}{6}$$

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

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

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

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

$$P(\mathit{Flight} = \mathit{on\ time}) = 0.6$$
$$P(\mathit{Flight} = \mathit{delayed}) = 0.3$$
$$P(\mathit{Flight} = \mathit{cancelled}) = 0.1$$

probability distribution

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

independence

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

independence

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

independence

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

independence

$P($$) = P($$)P($$)$

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

independence

$$P(\text{⚅⚃}) \neq P(\text{⚅})P(\text{⚃})$$

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

independence

$$P(\text{⚅⚃}) \neq P(\text{⚅})P(\text{⚃} \mid \text{⚅})$$

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

Bayes' Rule

$$P(a \wedge b) = P(b)\, P(a \mid b)$$

$$P(a \wedge b) = P(a)\, P(b \mid a)$$

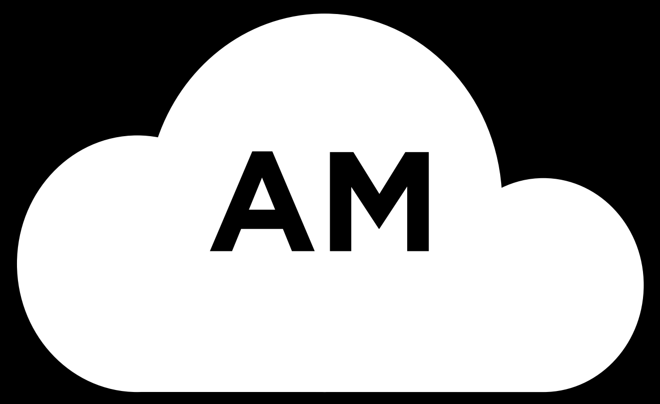
$$P(a)\,P(b \mid a) = P(b)\,P(a \mid b)$$

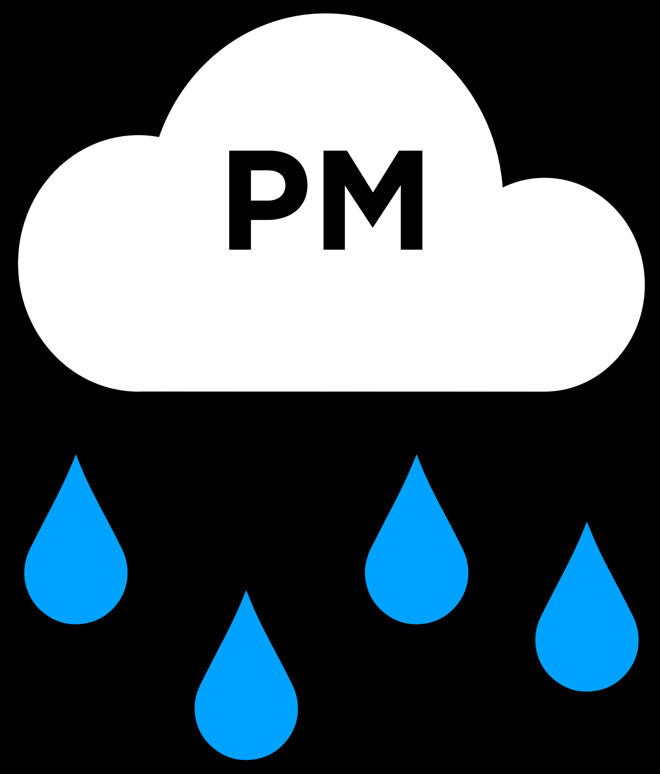
Bayes' Rule

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

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.

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

$$= \frac{(.8)(.1)}{.4}$$

$$= 0.2$$

Knowing

$$P(cloudy\ morning \mid rainy\ afternoon)$$

we can calculate

$$P(rainy\ afternoon \mid cloudy\ morning)$$

Knowing

$$P(\textit{visible effect} \mid \textit{unknown cause})$$

we can calculate

$$P(\textit{unknown cause} \mid \textit{visible effect})$$

Knowing

$$P(medical\ test\ result \mid disease)$$

we can calculate

$$P(disease \mid medical\ test\ result)$$

Knowing

$$P(blurry\ text \mid counterfeit\ bill)$$

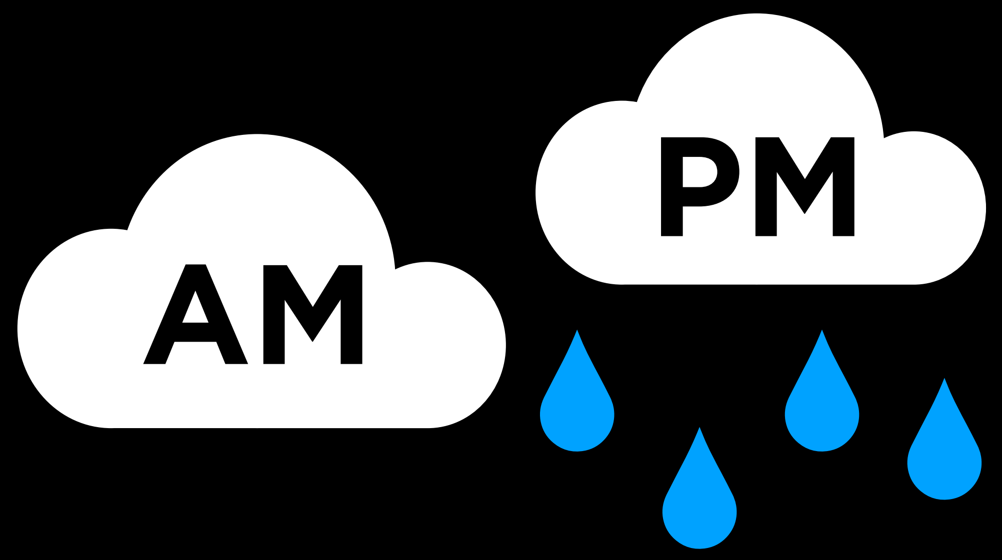
we can calculate

$$P(counterfeit\ bill \mid blurry\ text)$$

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

$$\mathbf{P}(C \mid rain)$$

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

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

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

Probability Rules

Negation

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

Inclusion-Exclusion

$$P(a \vee b) = P(a) + P(b) - P(a \wedge b)$$

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

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

$$P(C = cloud)$$

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

$$= 0.08 + 0.32$$

$$= 0.40$$

Conditioning

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

Conditioning

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

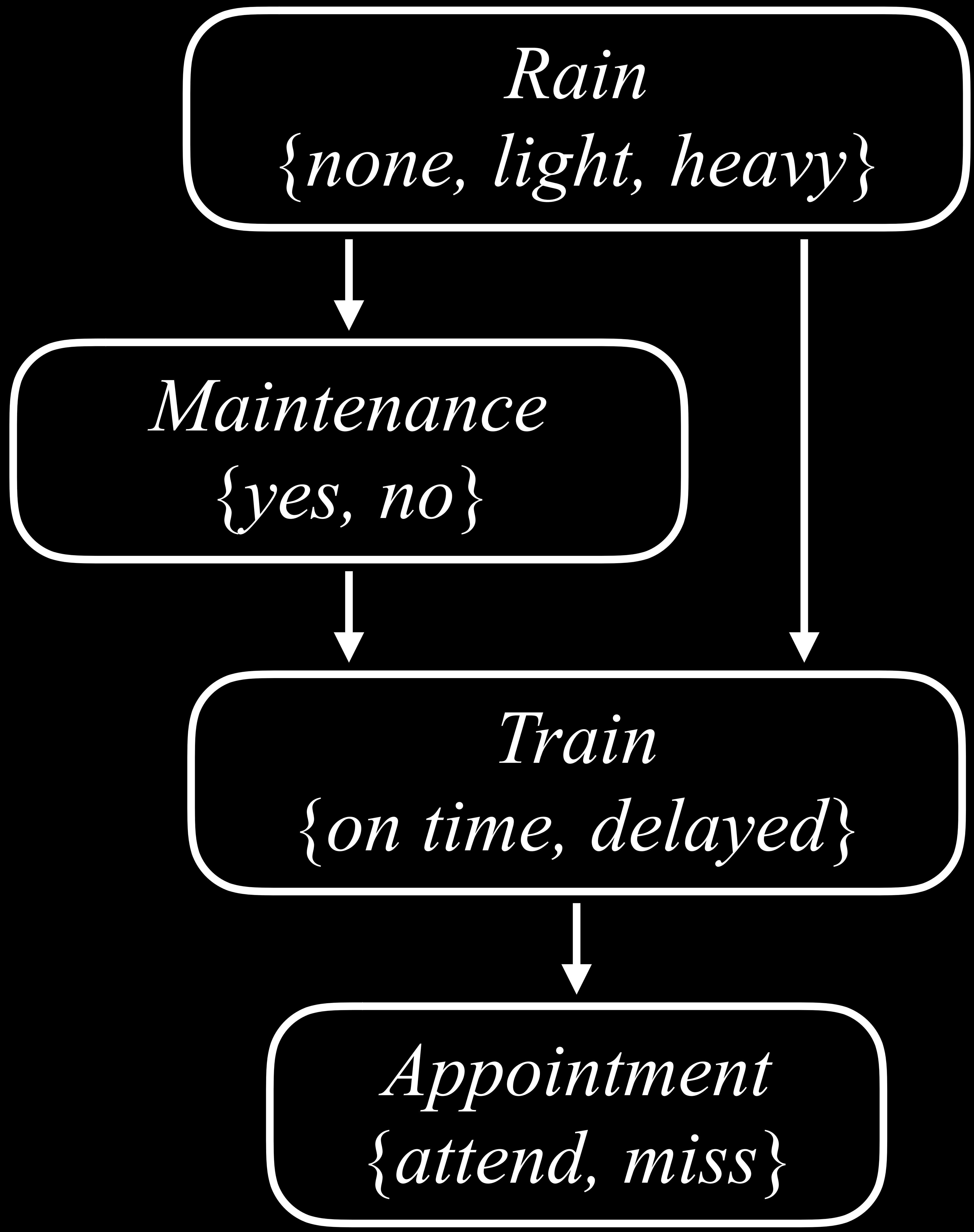
Bayesian Networks

Bayesian network

data structure that represents the dependencies among random variables

Bayesian network

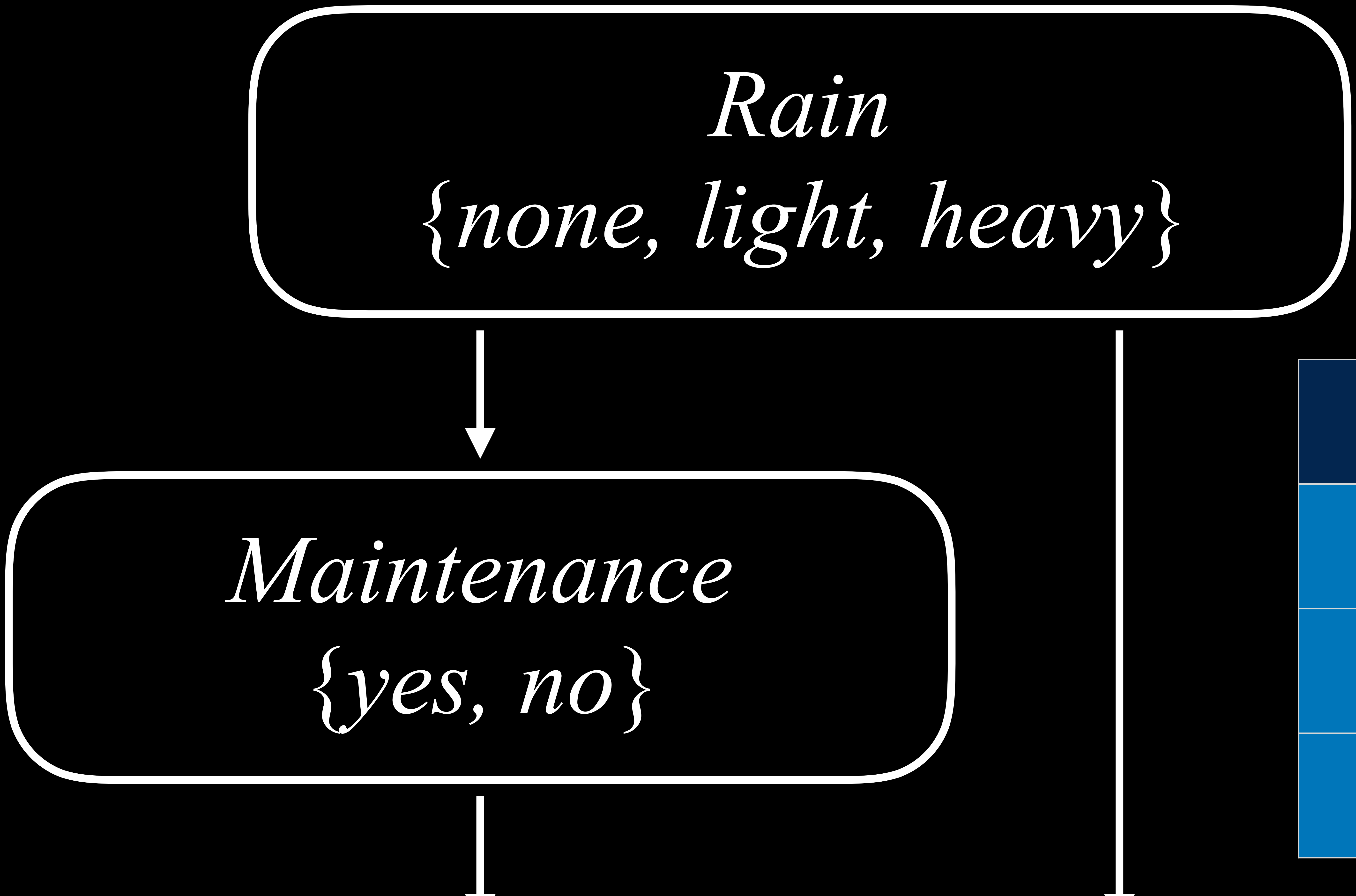
- directed graph
- each node represents a random variable
- arrow from X to Y means X is a parent of Y
- each node X has probability distribution $\mathbf{P}(X \mid Parents(X))$


Rain
{none, light, heavy}
Maintenance
{yes, no}
Train
{on time, delayed}
Appointment
{attend, miss}

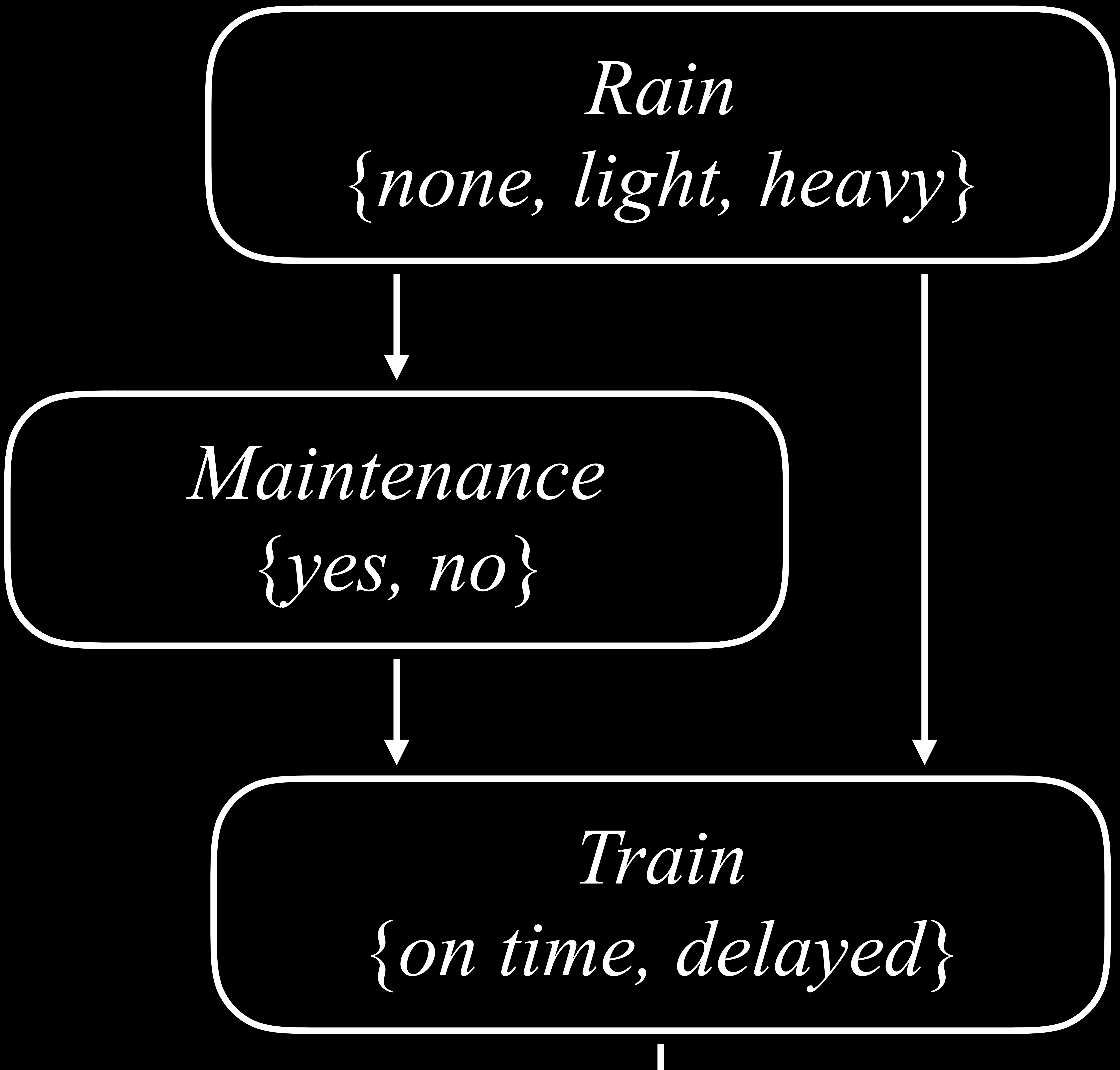
Rain

$\{none, light, heavy\}$

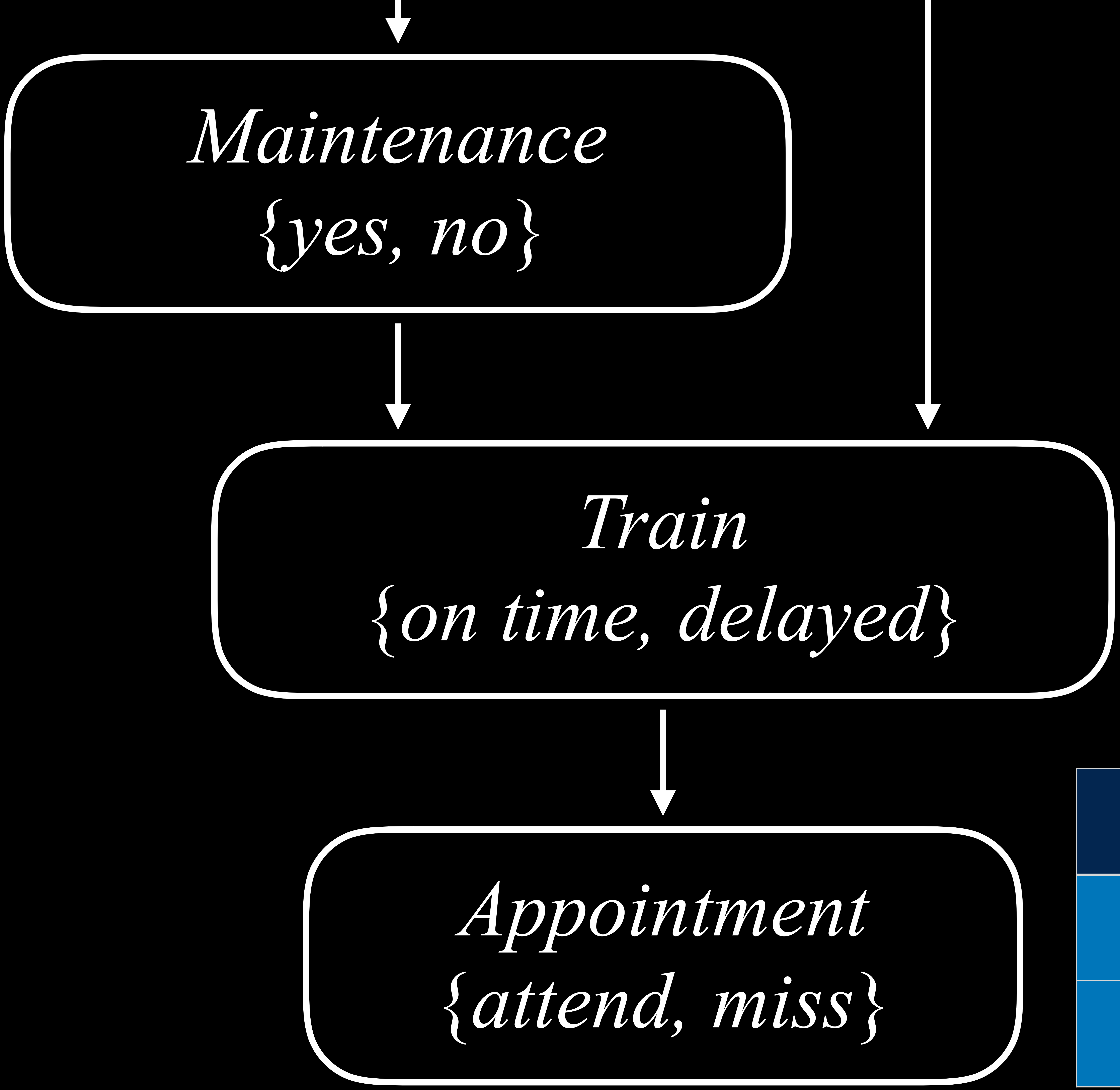
none	*light*	*heavy*
0.7	0.2	0.1



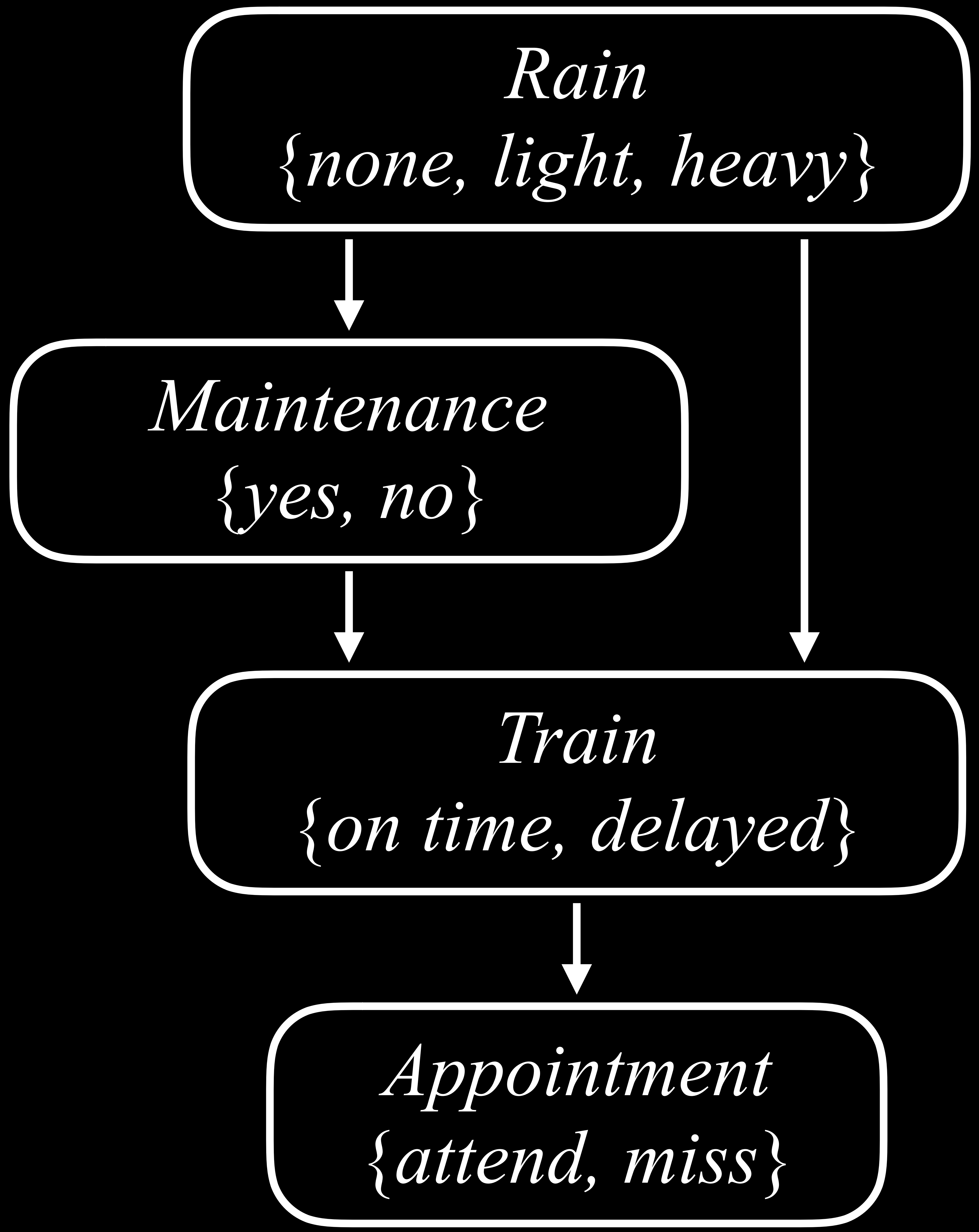
R	*yes*	*no*
none	0.4	0.6
light	0.2	0.8
heavy	0.1	0.9



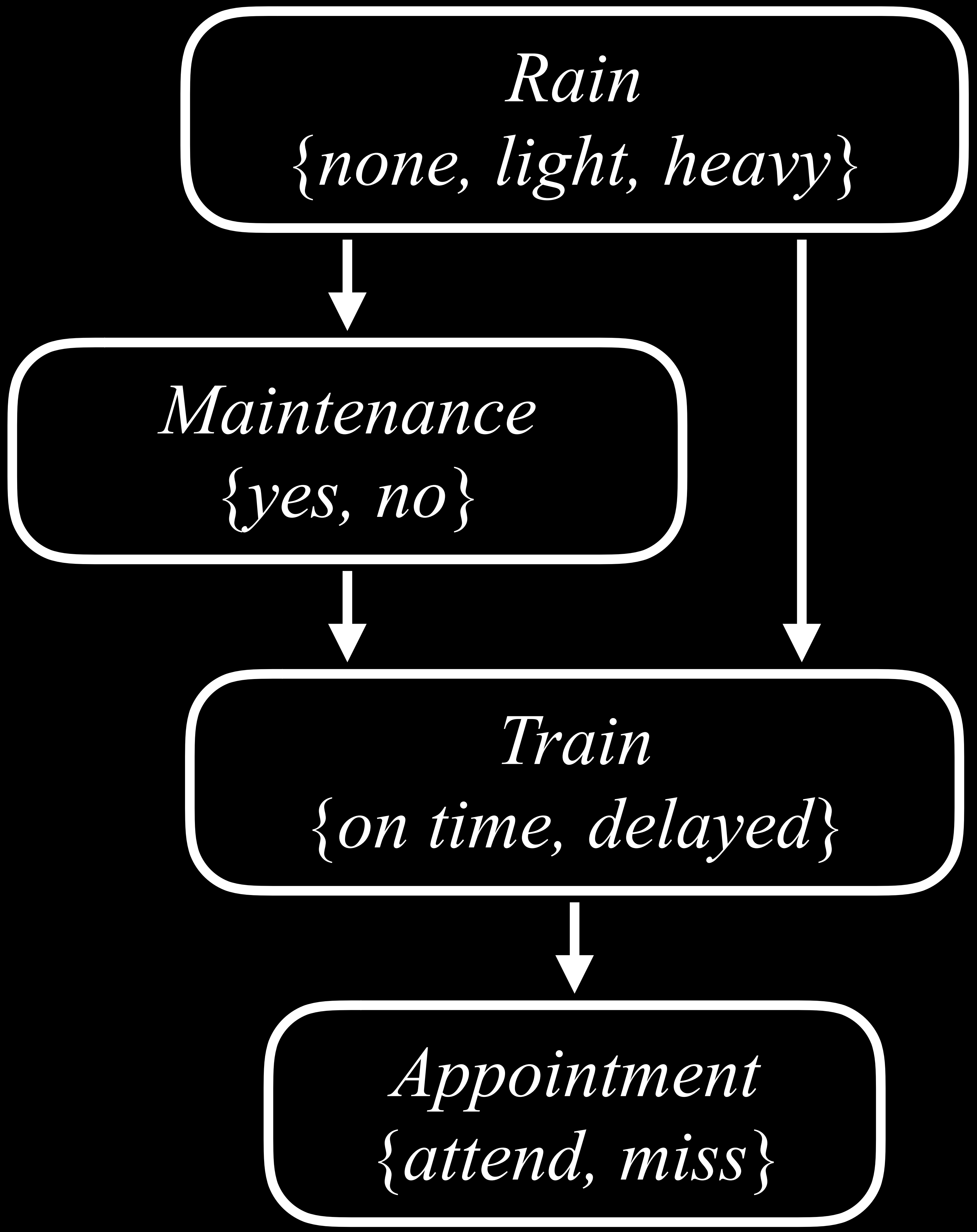
R	*M*	*on time*	*delayed*
none	*yes*	0.8	0.2
none	*no*	0.9	0.1
light	*yes*	0.6	0.4
light	*no*	0.7	0.3
heavy	*yes*	0.4	0.6
heavy	*no*	0.5	0.5



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}
Appointment
{attend, miss}

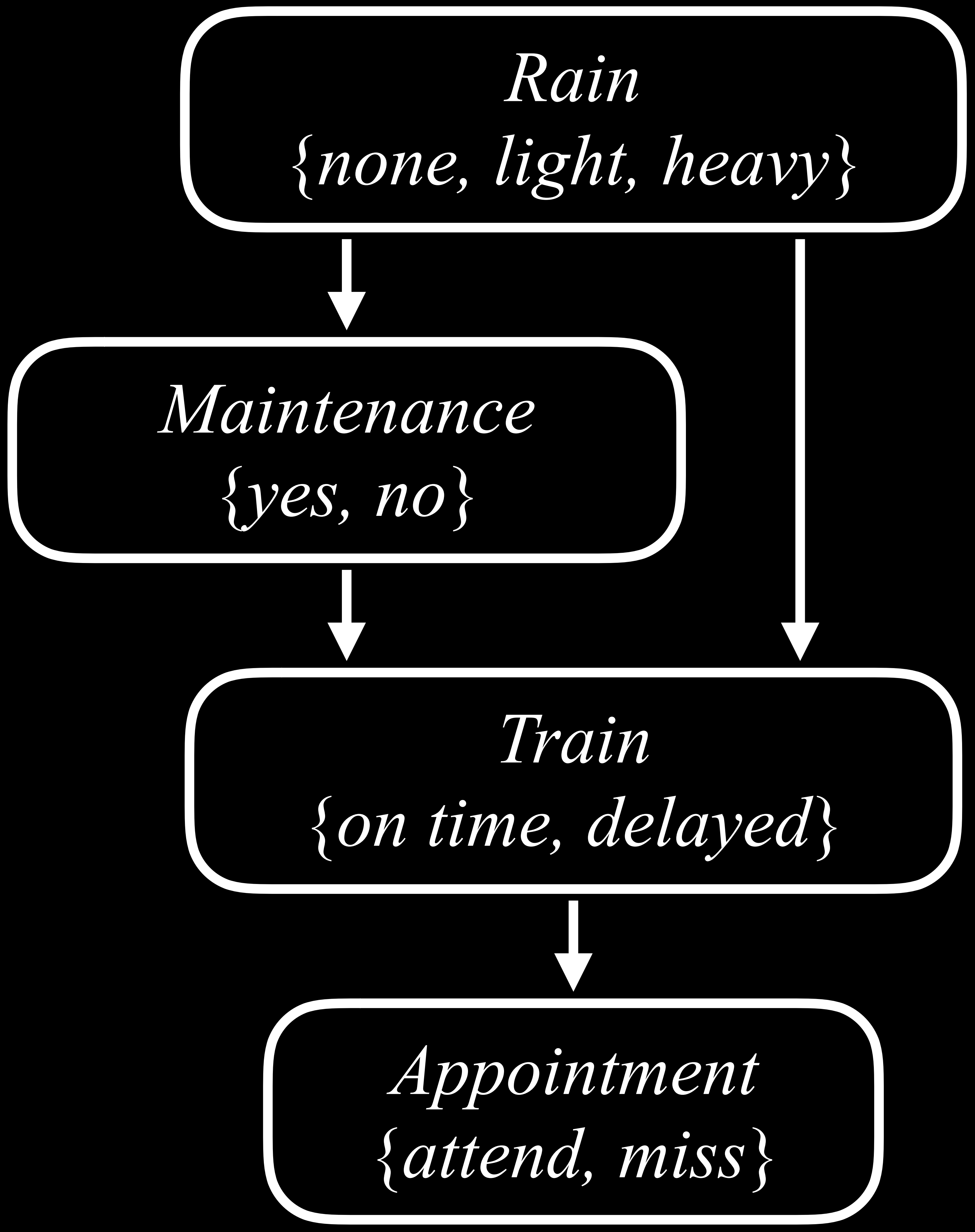
Computing Joint Probabilities



$$P(light)$$

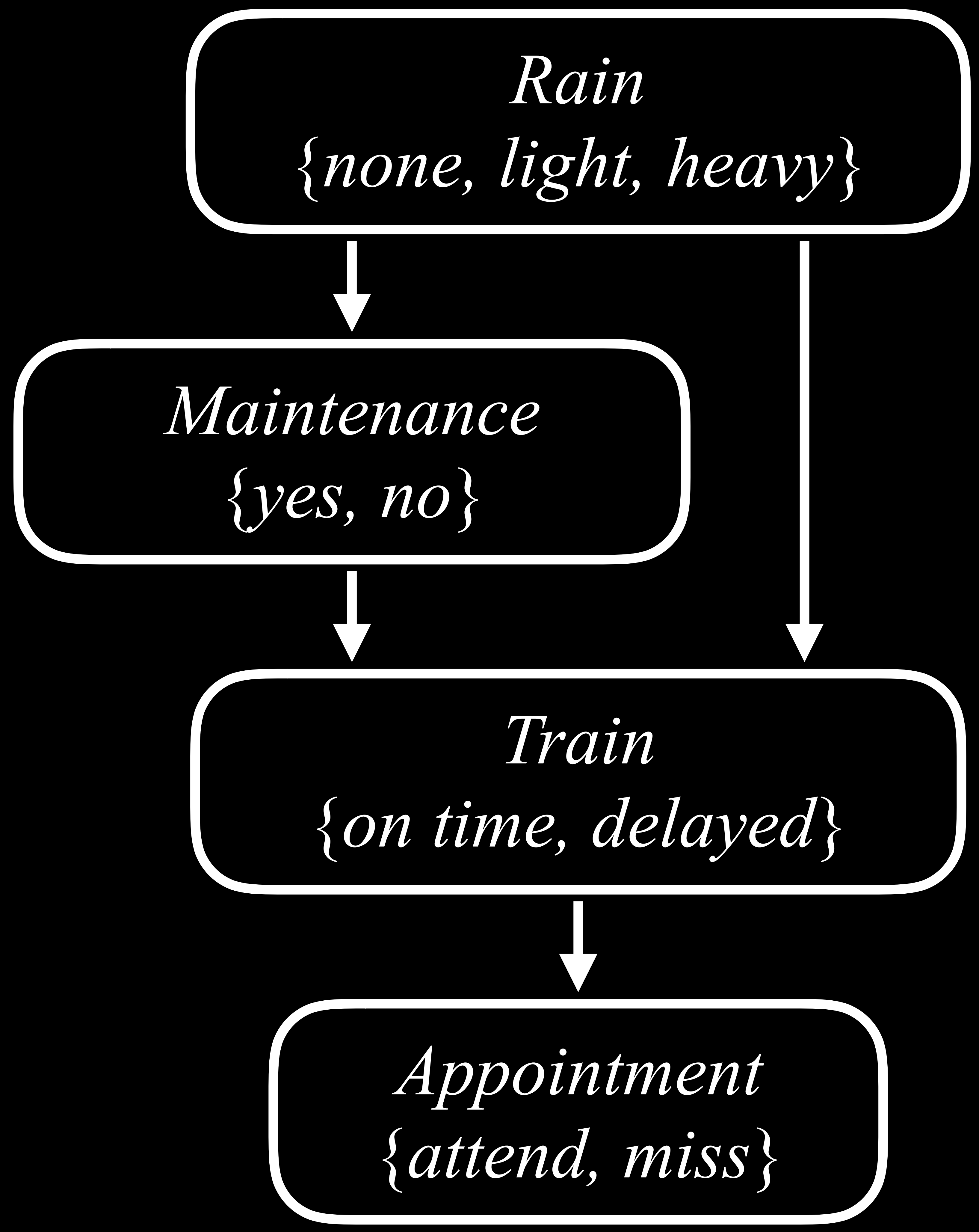
$$P(light)$$

Computing Joint Probabilities



$$P(light, no)$$

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

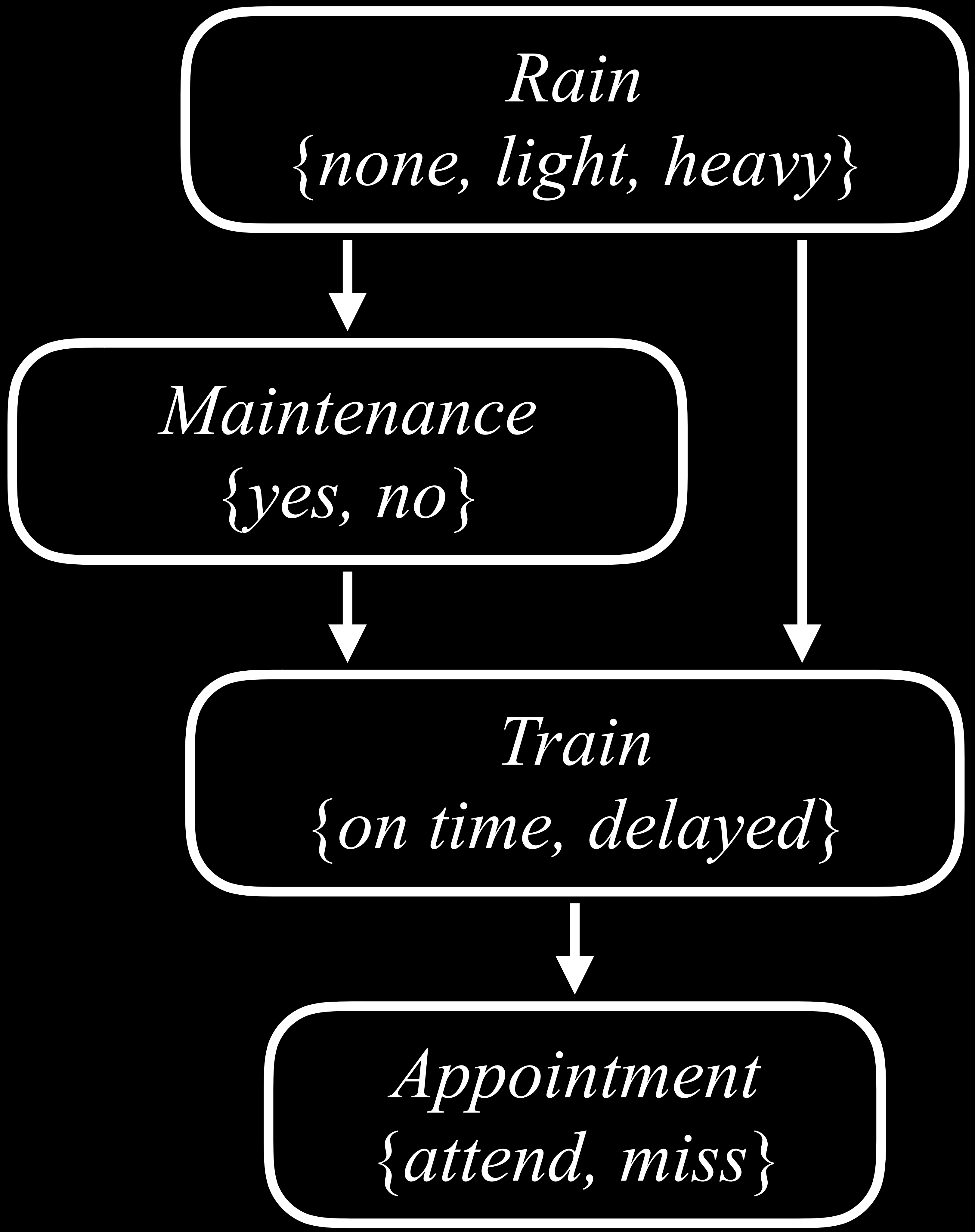


Computing Joint Probabilities

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

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

Computing Joint Probabilities



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

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

Inference

Inference

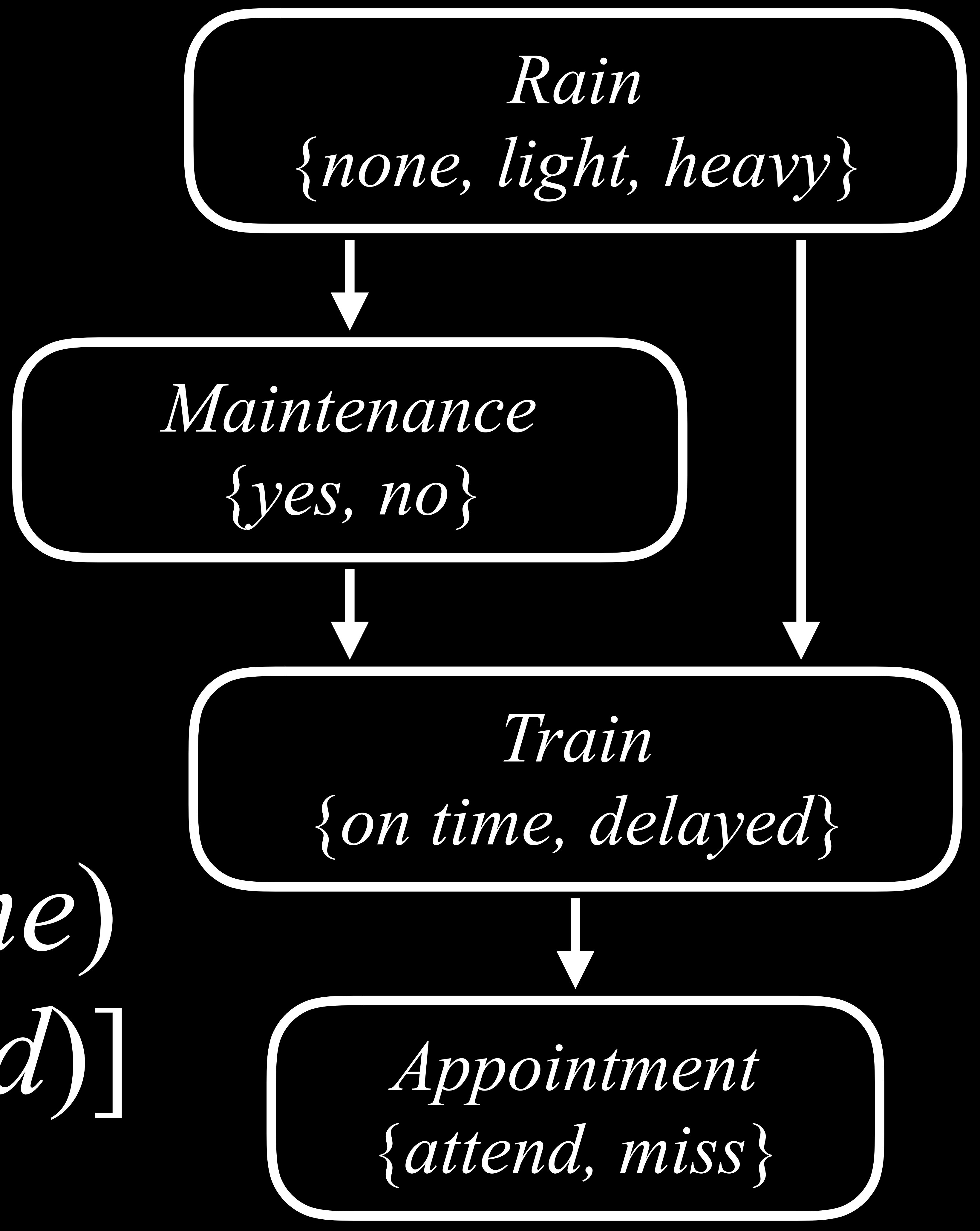
- Query $\mathbf{X}$: variable for which to compute distribution
- Evidence variables $\mathbf{E}$: observed variables for event $\mathbf{e}$
- Hidden variables $\mathbf{Y}$: non-evidence, non-query variable.

- Goal: Calculate $\mathbf{P}(\mathbf{X} \mid \mathbf{e})$

$P(\text{Appointment} \mid \textit{light}, \textit{no})$

$= \alpha\, P(\text{Appointment}, \textit{light}, \textit{no})$

$= \alpha\, [P(\text{Appointment}, \textit{light}, \textit{no}, \textit{on time}) + P(\text{Appointment}, \textit{light}, \textit{no}, \textit{delayed})]$



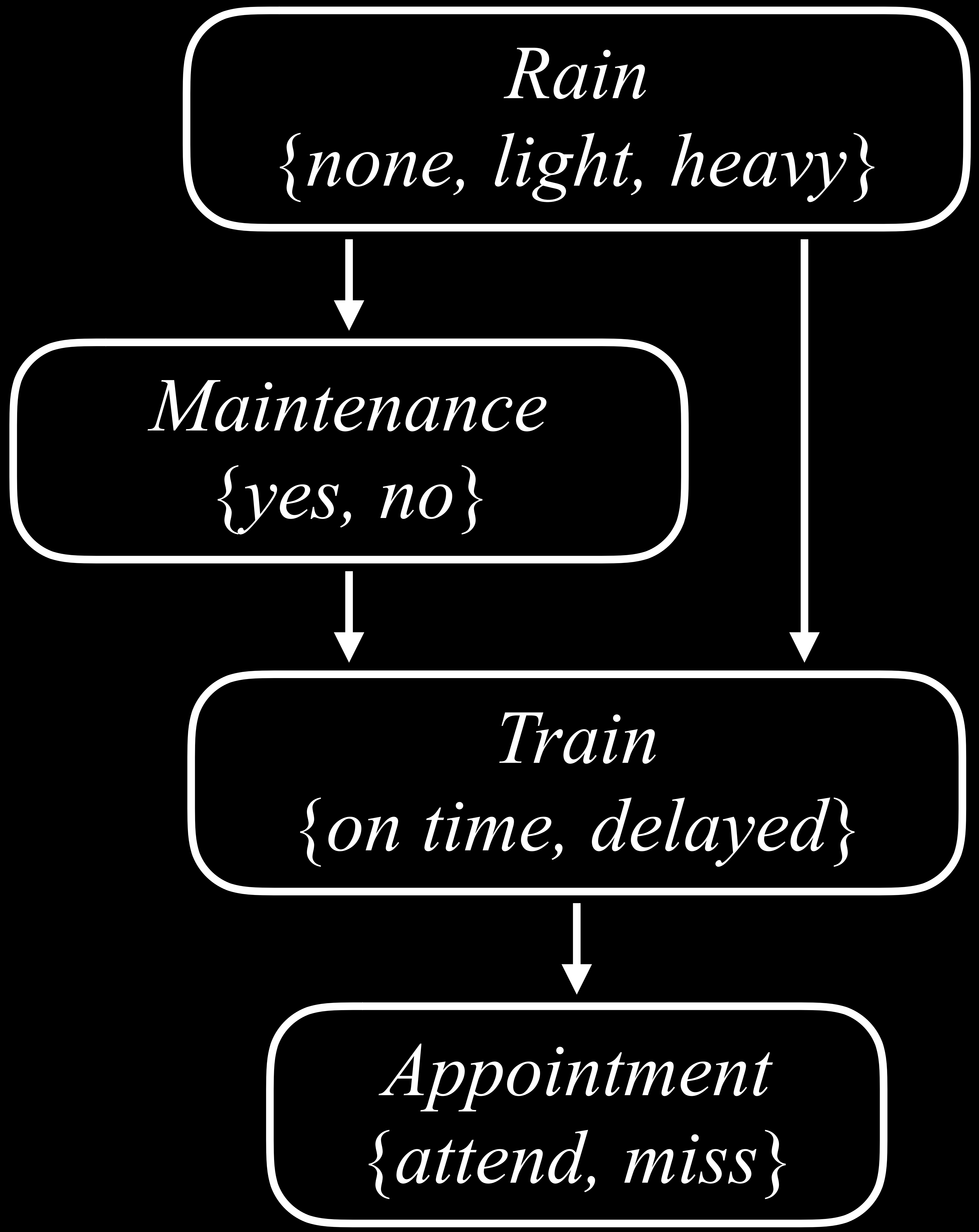
Inference by Enumeration

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

X is the query variable.
$\mathbf{e}$ is the evidence.
$\mathbf{y}$ ranges over values of hidden variables.
α normalizes the result.

Approximate Inference

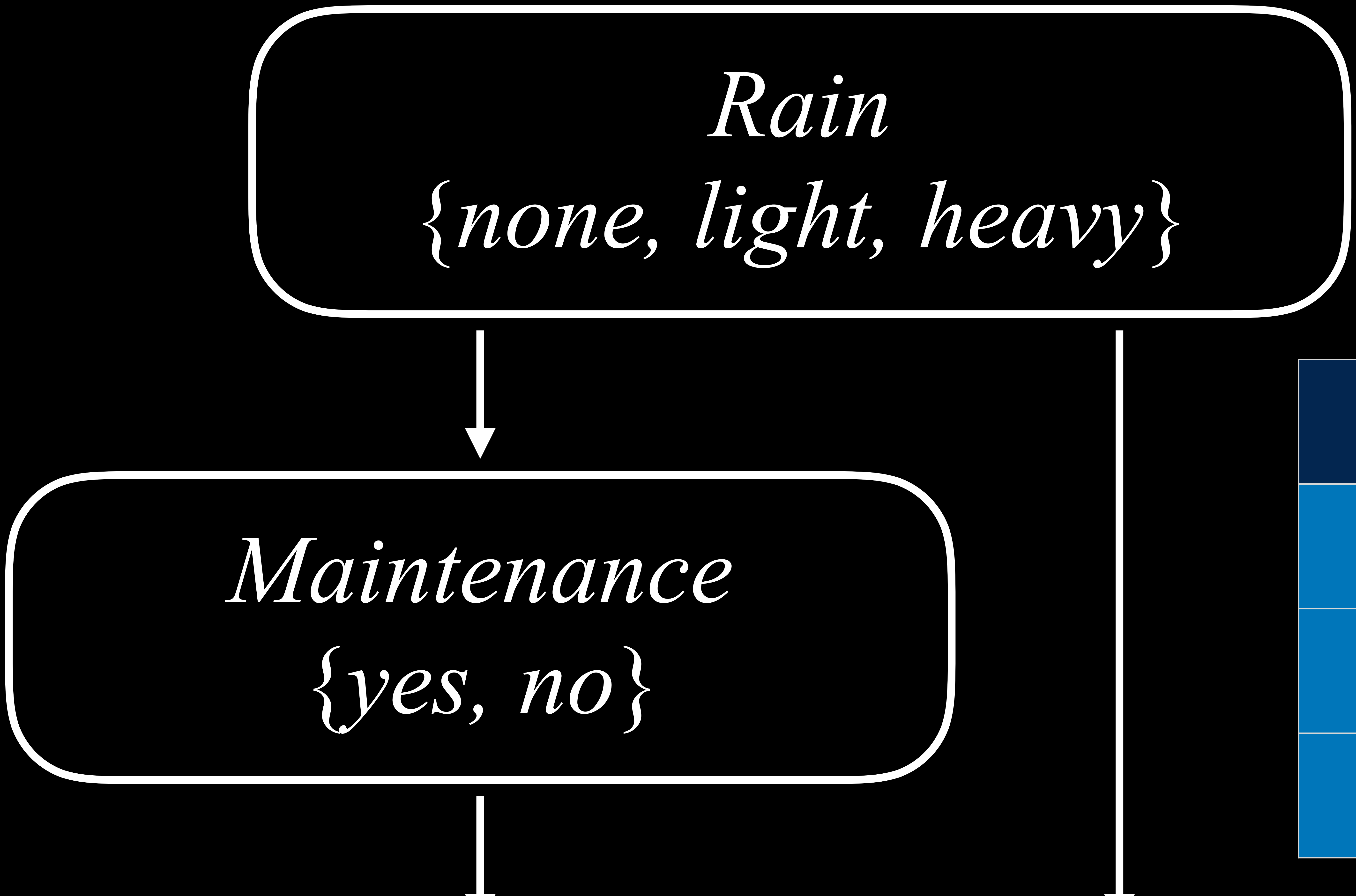
Sampling


Rain
{none, light, heavy}
Maintenance
{yes, no}
Train
{on time, delayed}
Appointment
{attend, miss}

R = *none*

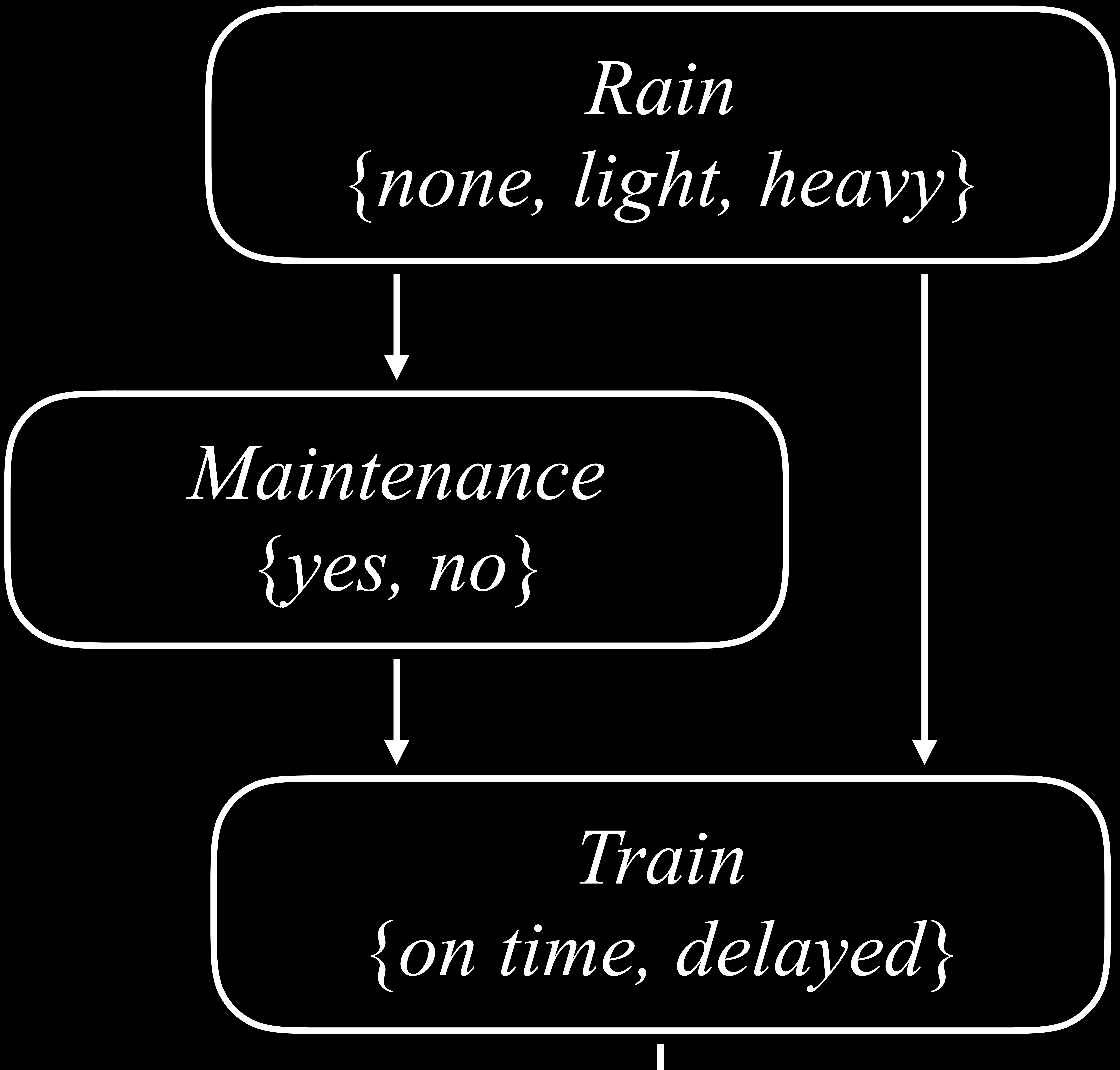
Rain
{*none*, *light*, *heavy*}

none	*light*	*heavy*
0.7	0.2	0.1



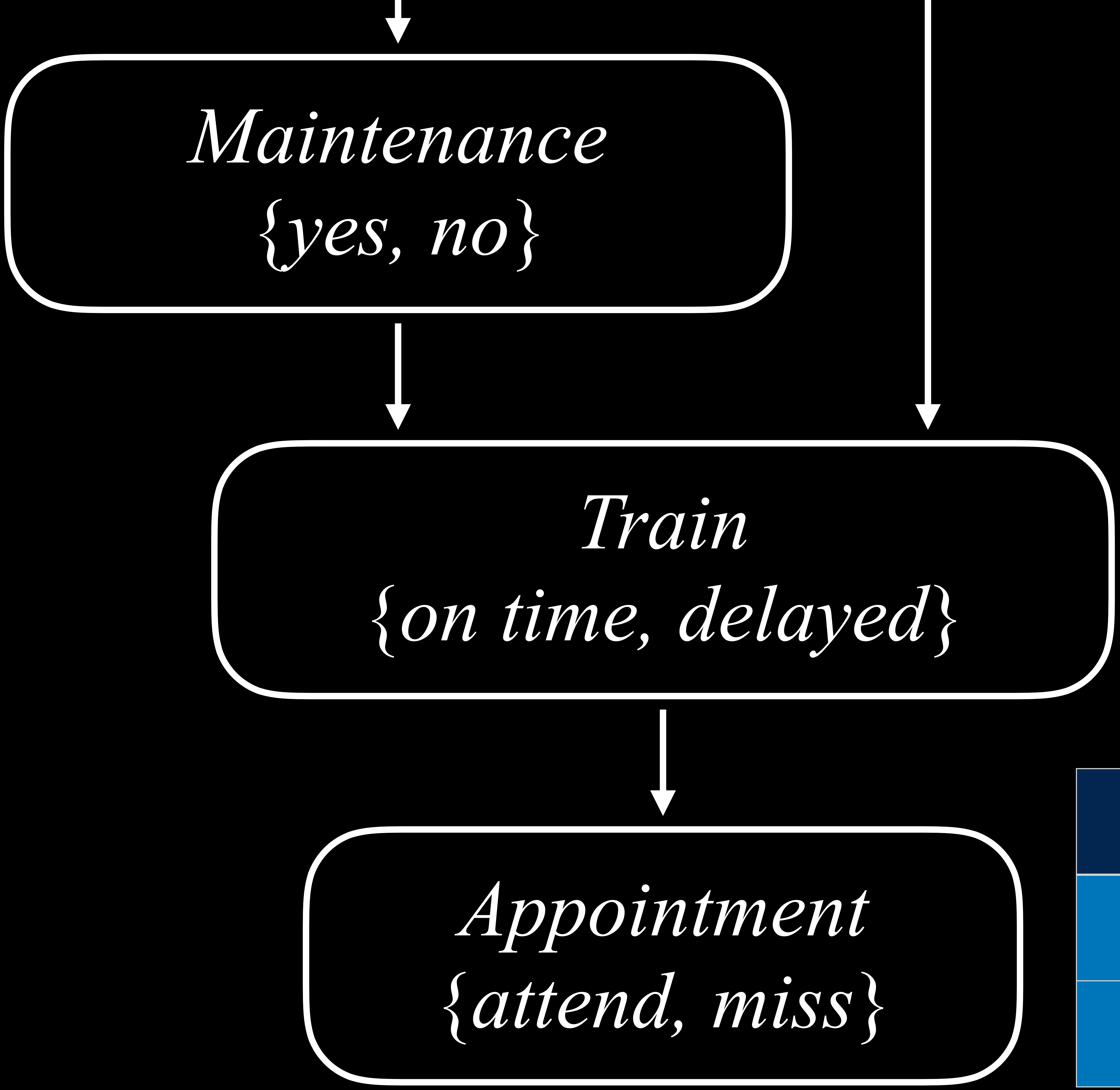
R = *none*
M = *yes*

R	*yes*	*no*
none	0.4	0.6
light	0.2	0.8
heavy	0.1	0.9



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



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

T	*attend*	*miss*
on time	0.9	0.1
delayed	0.6	0.4

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

R = *light*	R = *light*	R = *none*	R = *none*
M = *no*	M = *yes*	M = *no*	M = *yes*
T = *on time*	T = *delayed*	T = *on time*	T = *on time*
A = *miss*	A = *attend*	A = *attend*	A = *attend*

R = *none*	R = *none*	R = *heavy*	R = *light*
M = *yes*	M = *yes*	M = *no*	M = *no*
T = *on time*	T = *on time*	T = *delayed*	T = *on time*
A = *attend*	A = *attend*	A = *miss*	A = *attend*

$$P(\mathit{Train} = \mathit{on\ time})\ ?$$

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*

R = *light*	R = *light*	R = *none*	R = *none*
M = *no*	M = *yes*	M = *no*	M = *yes*
T = *on time*	T = *delayed*	T = *on time*	T = *on time*
A = *miss*	A = *attend*	A = *attend*	A = *attend*

R = *none*	R = *none*	R = *heavy*	R = *light*
M = *yes*	M = *yes*	M = *no*	M = *no*
T = *on time*	T = *on time*	T = *delayed*	T = *on time*
A = *attend*	A = *attend*	A = *miss*	A = *attend*

$$P(\text{Rain} = \textit{light} \mid \text{Train} = \textit{on time})\ ?$$

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*

R = *light*	R = *light*	R = *none*	R = *none*
M = *no*	M = *yes*	M = *no*	M = *yes*
T = *on time*	T = *delayed*	T = *on time*	T = *on time*
A = *miss*	A = *attend*	A = *attend*	A = *attend*

R = *none*	R = *none*	R = *heavy*	R = *light*
M = *yes*	M = *yes*	M = *no*	M = *no*
T = *on time*	T = *on time*	T = *delayed*	T = *on time*
A = *attend*	A = *attend*	A = *miss*	A = *attend*

$R = light$	$R = light$	$R = none$	$R = none$	$R = none$	$R = none$	$R = heavy$	$R = light$
$M = no$	$M = yes$	$M = no$	$M = yes$	$M = yes$	$M = yes$	$M = no$	$M = no$
$T = on\ time$	$T = delayed$	$T = on\ time$	$T = on\ time$	$T = on\ time$	$T = on\ time$	$T = delayed$	$T = on\ time$
$A = miss$	$A = attend$	$A = attend$	$A = attend$	$A = attend$	$A = attend$	$A = miss$	$A = attend$

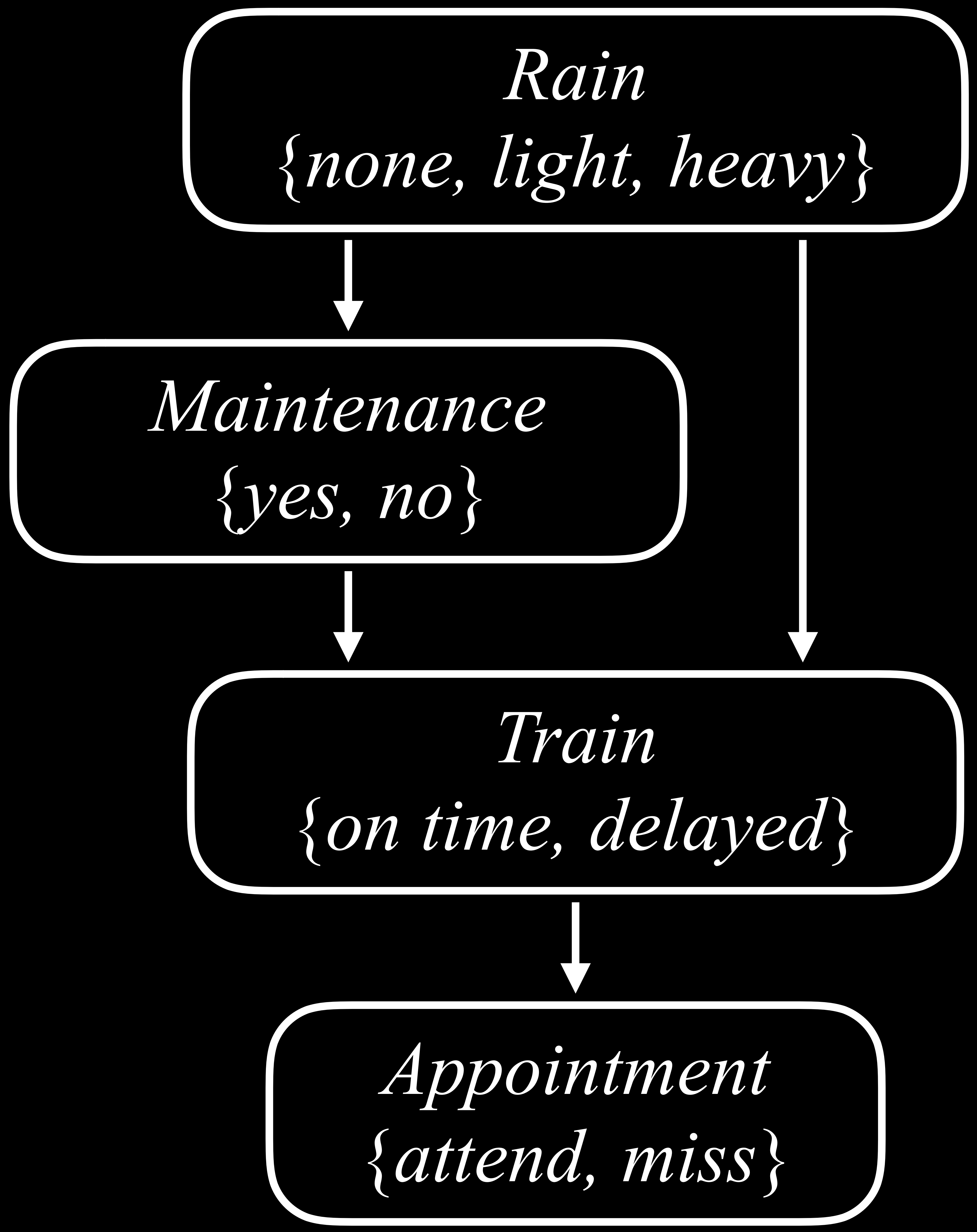
Rejection Sampling

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.

$$P(\text{Rain} = \textit{light} \mid \text{Train} = \textit{on time})\ ?$$

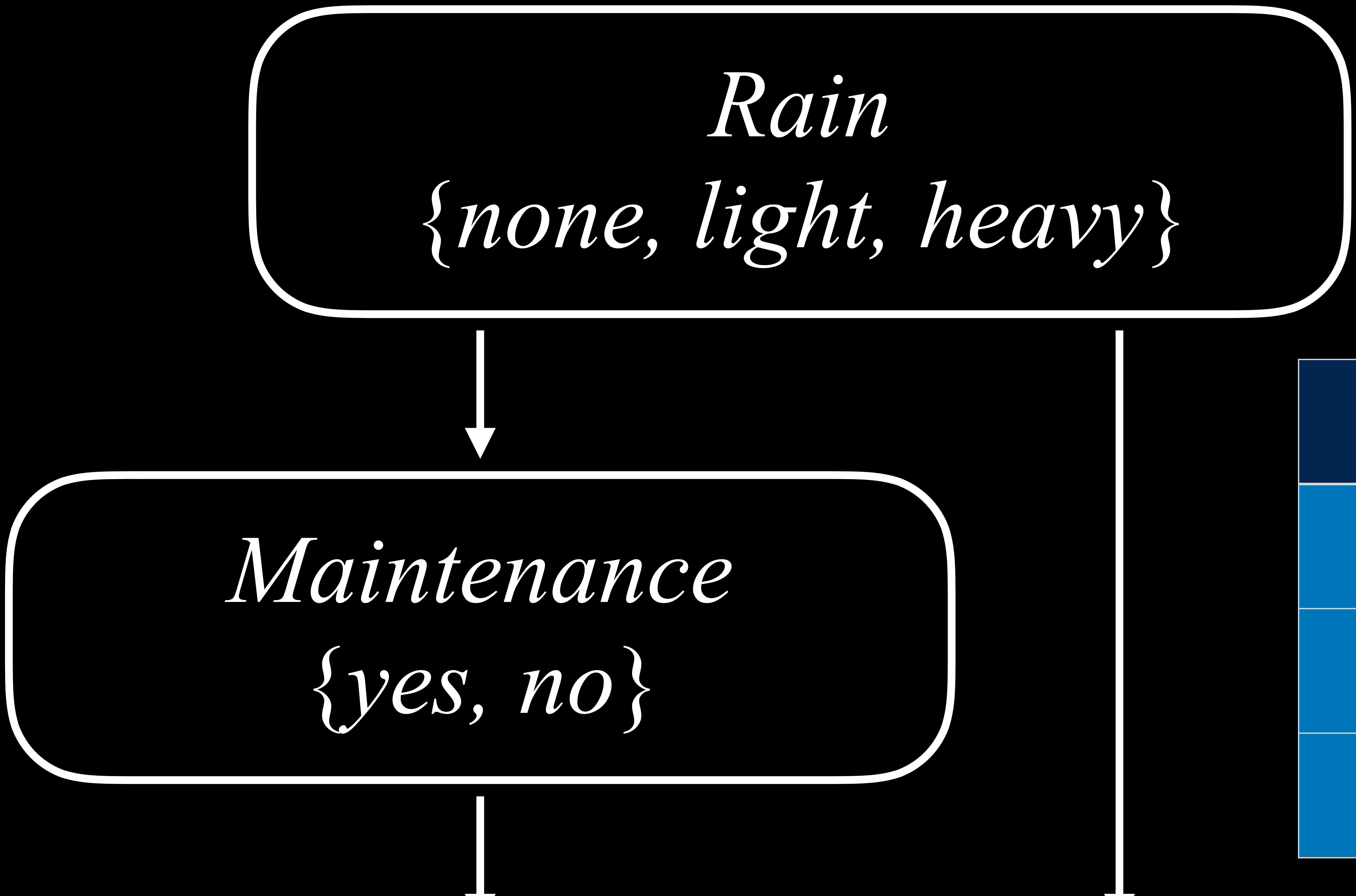

Rain
{none, light, heavy}
Maintenance
{yes, no}
Train
{on time, delayed}
Appointment
{attend, miss}

R = *light*
T = *on time*

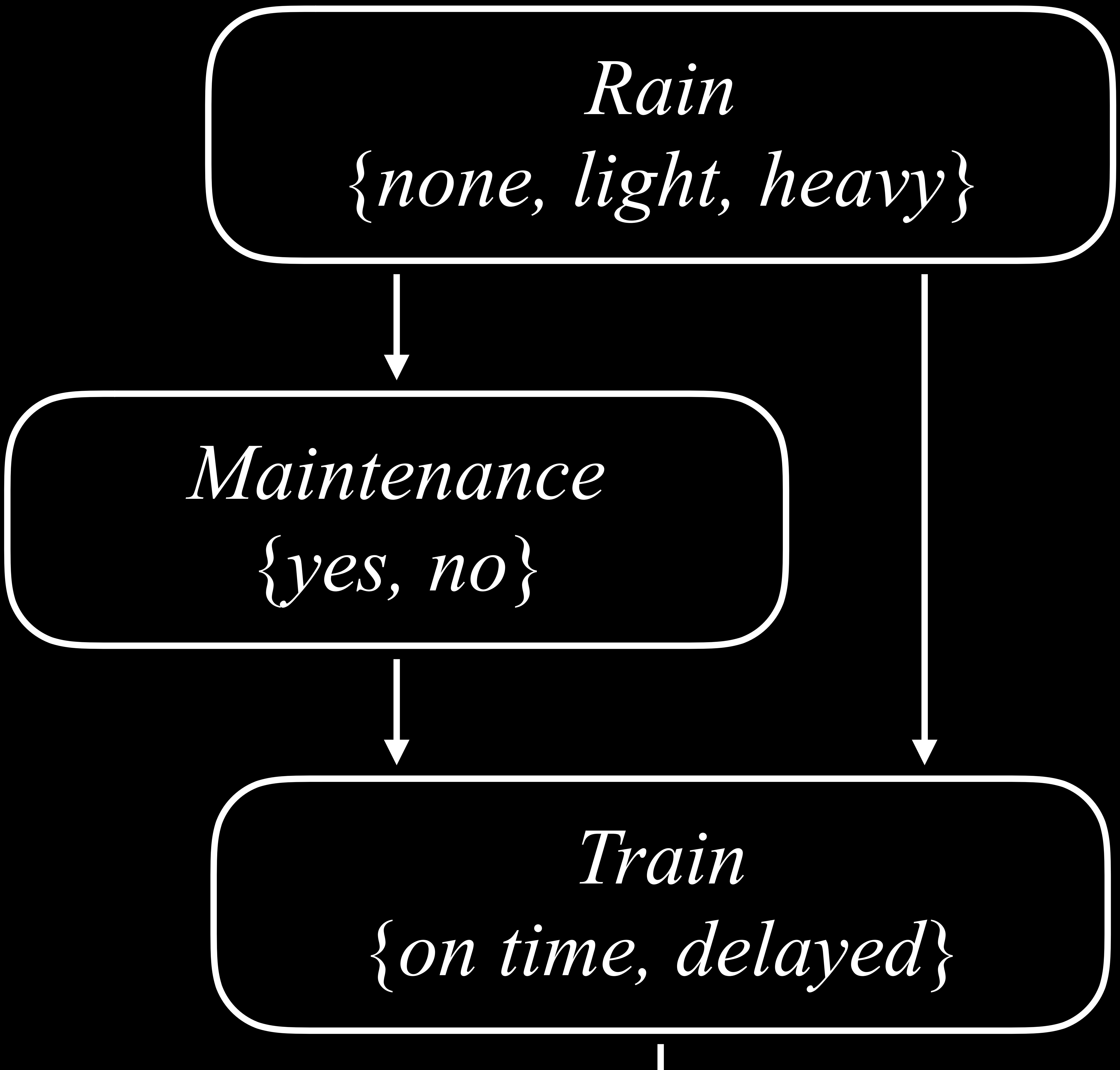
Rain
{*none*, *light*, *heavy*}

none	*light*	*heavy*
0.7	0.2	0.1

R = *light*
M = *yes*
T = *on time*

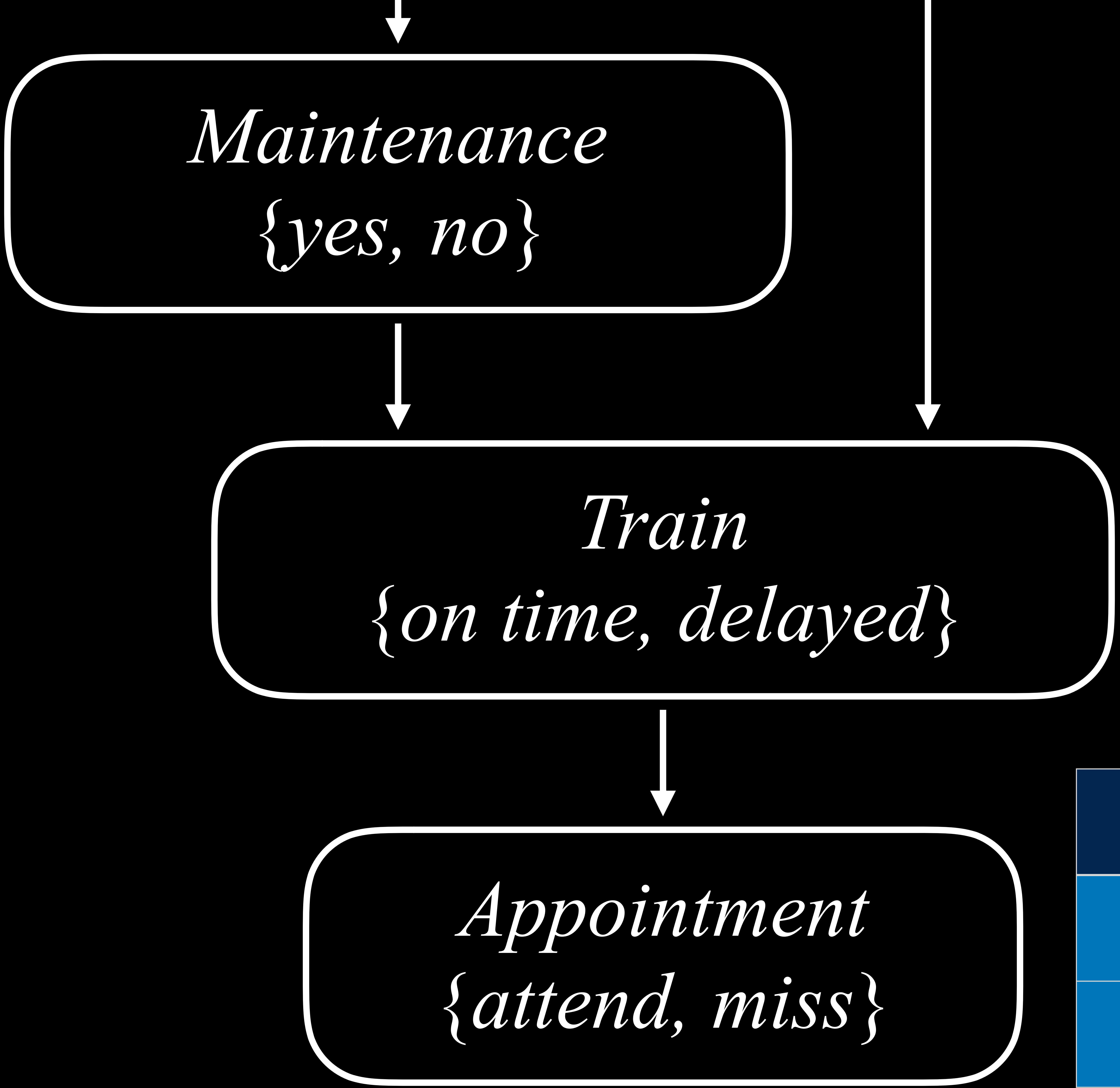


R	*yes*	*no*
none	0.4	0.6
light	0.2	0.8
heavy	0.1	0.9



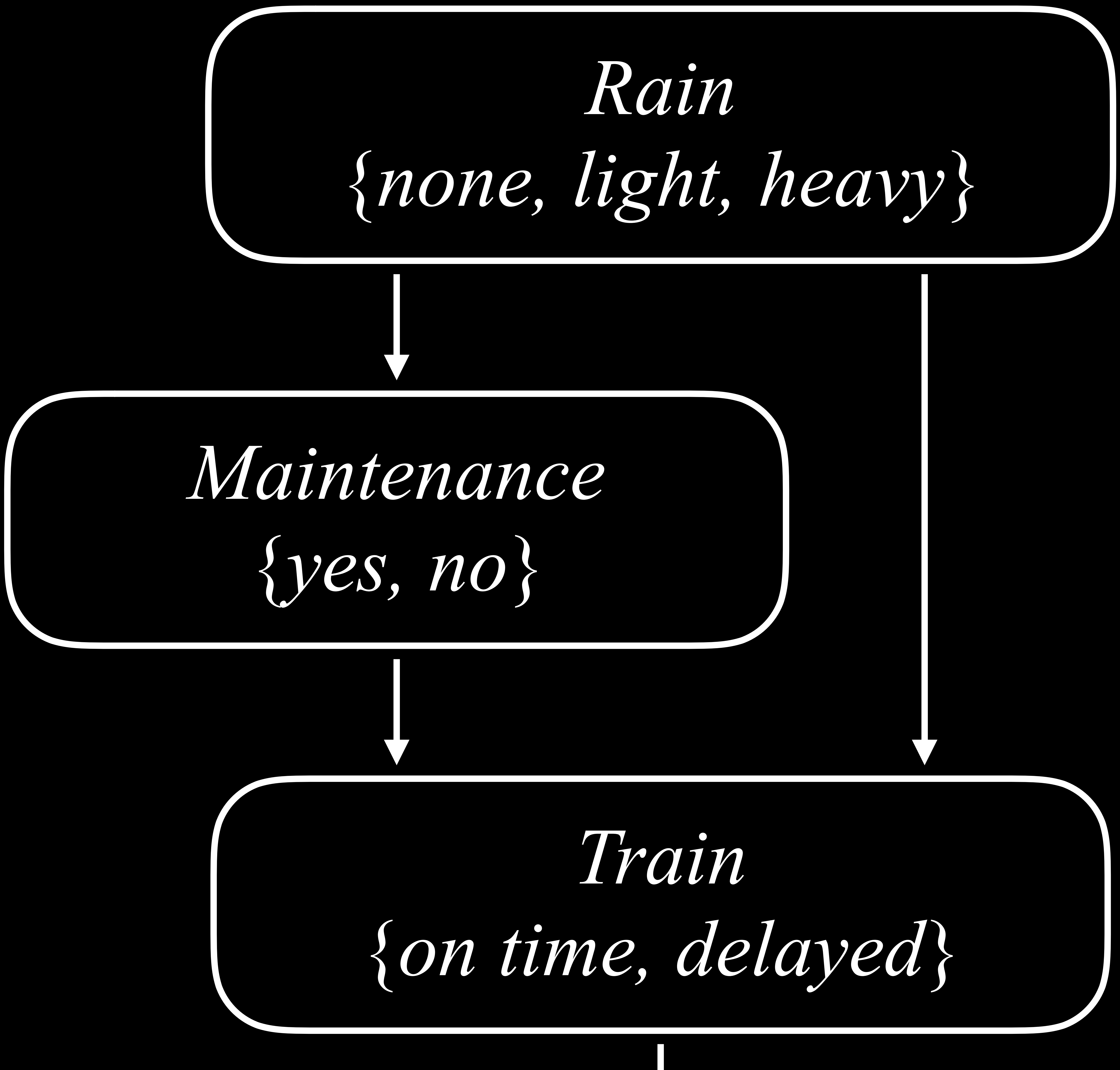
R = *light*
M = *yes*
T = *on time*

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



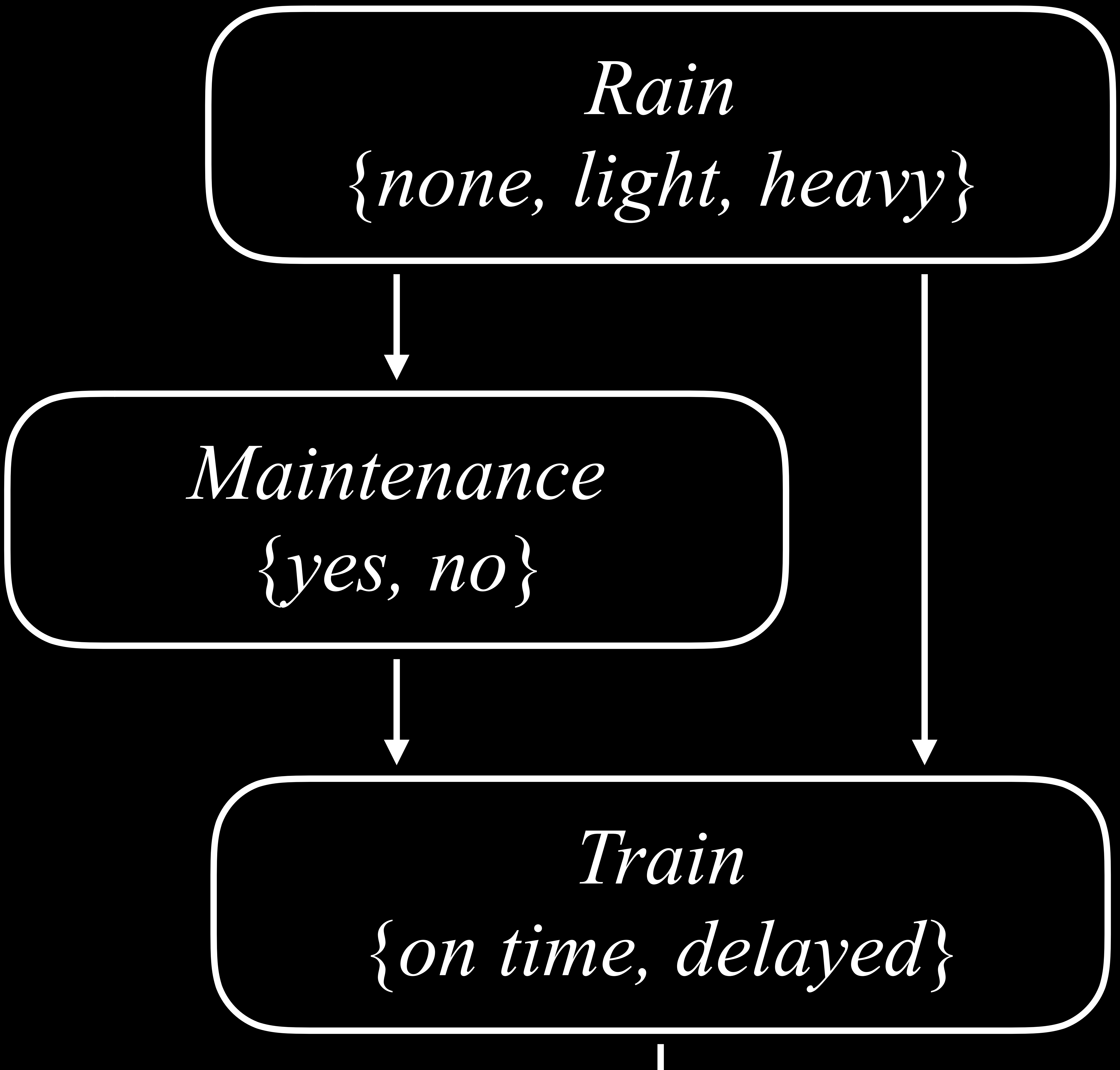
R = *light*
M = *yes*
T = *on time*
A = *attend*

T	*attend*	*miss*
on time	0.9	0.1
delayed	0.6	0.4



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

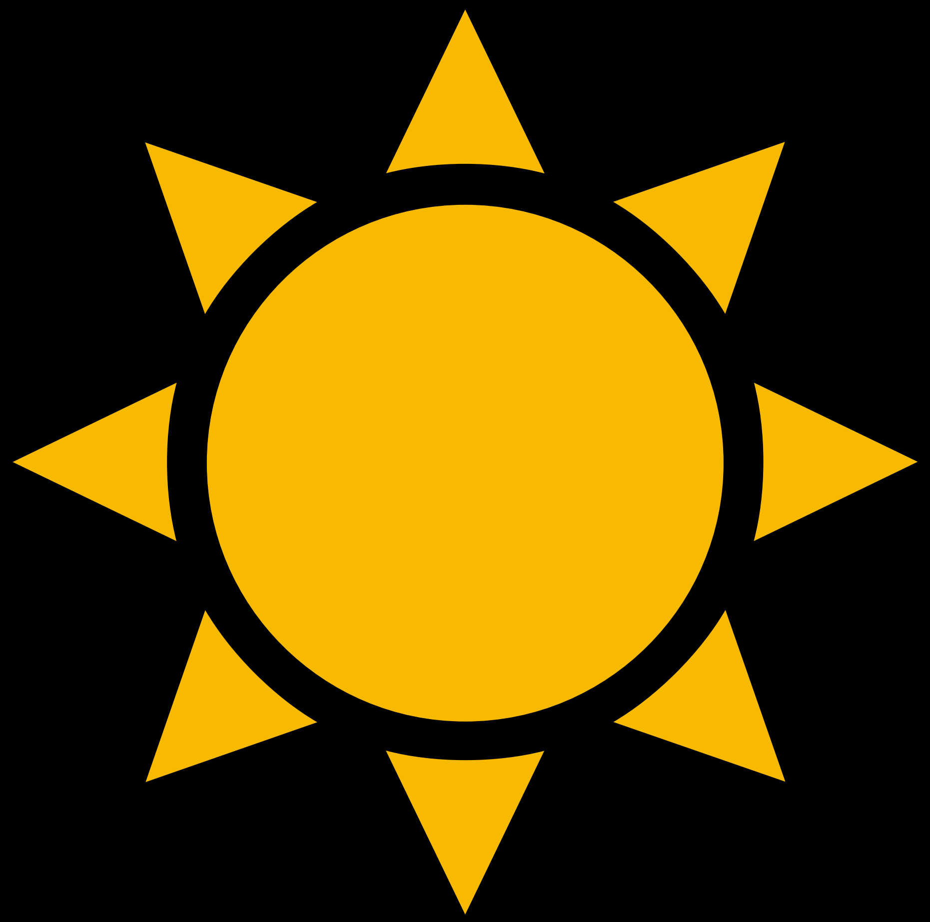
R	*M*	*on time*	*delayed*
none	*yes*	0.8	0.2
none	*no*	0.9	0.1
light	*yes*	0.6	0.4
light	*no*	0.7	0.3
heavy	*yes*	0.4	0.6
heavy	*no*	0.5	0.5



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



X_t: Weather at time t

Markov assumption

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

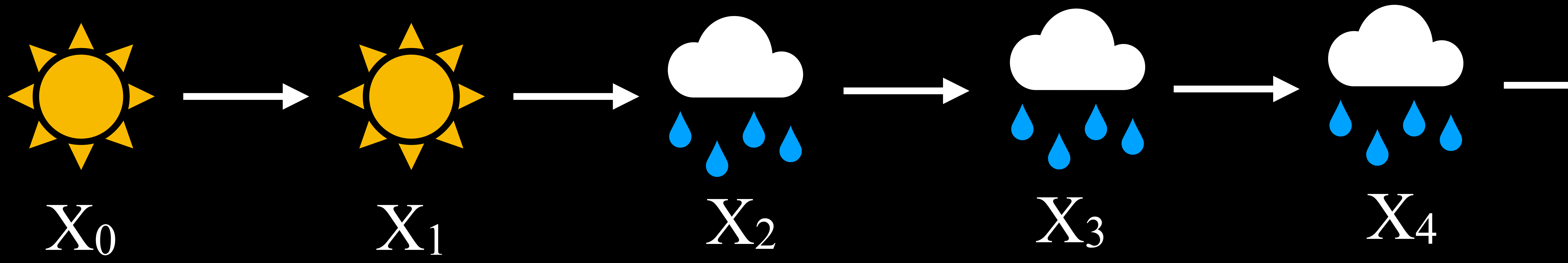
Markov Chain

Markov chain

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

Transition Model

		Tomorrow (X_{t+1})	
			
Today (X_t)		0.8	0.2
		0.3	0.7


X_0
X_1
X_2
X_3
X_4

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 Model

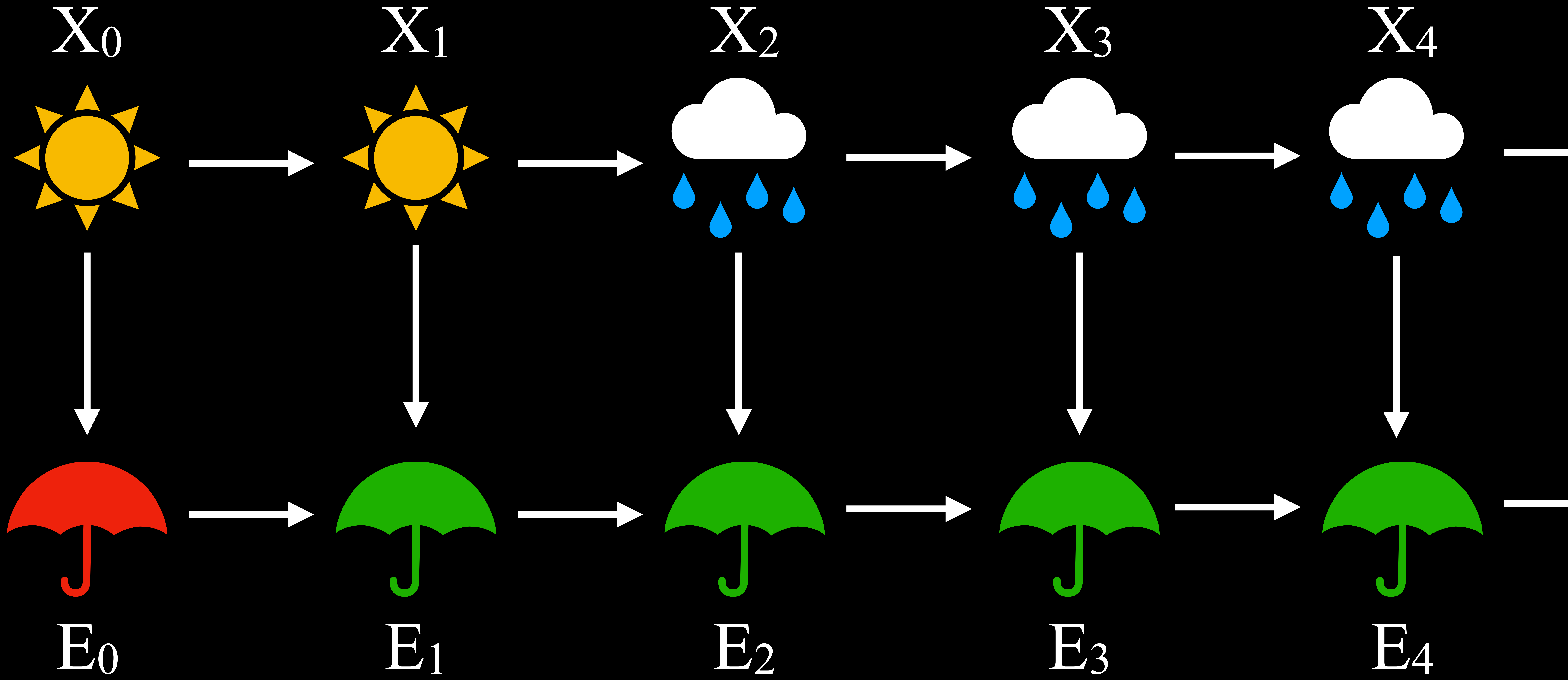
Observation (E_t)

State (X_t)

		
	0.2	0.8
	0.9	0.1

sensor Markov assumption

the assumption that the evidence variable depends only the corresponding state


X0
X1
X2
X3
X4
E0
E1
E2
E3
E4

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

Introduction to

Artificial Intelligence

with Python