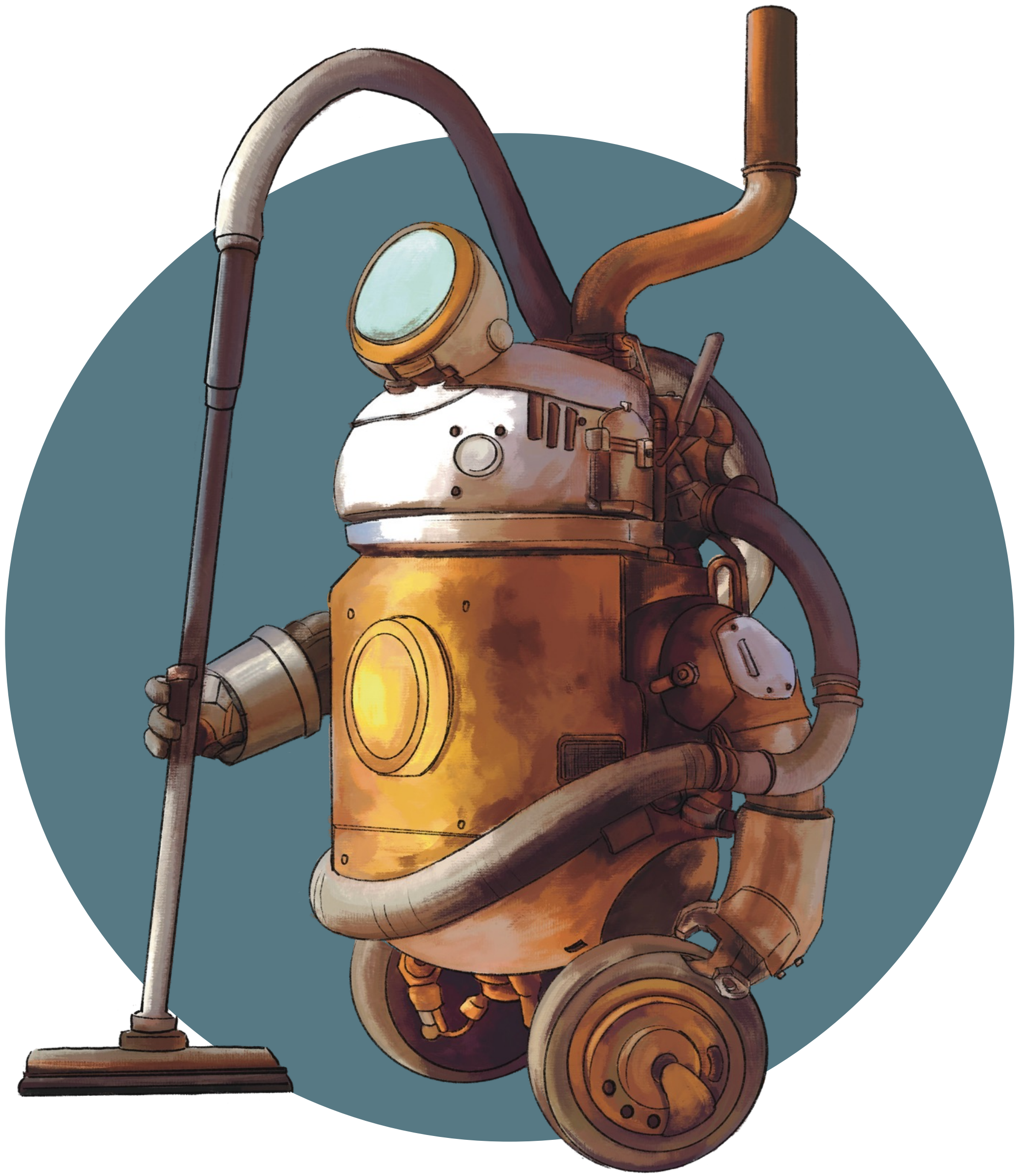
Lecture 8:
Actions & Markov chains



We control the actions,
not their outcomes.

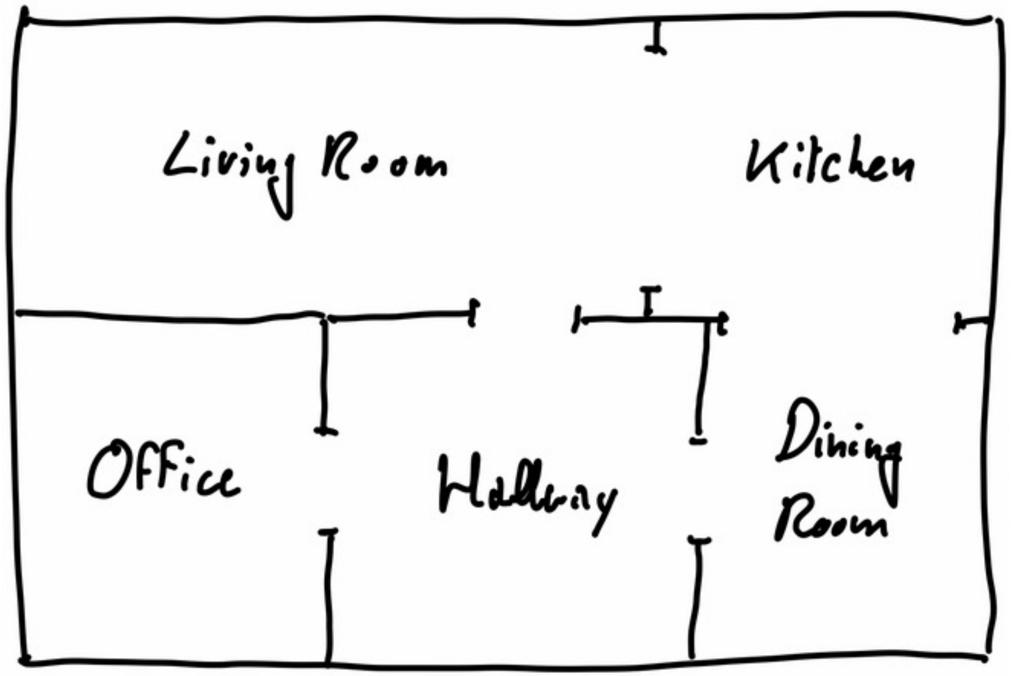
Frank Dellaert • Seth Hutchinson

After two right commands, where could the robot be?

Model one action → update our belief → repeat over time

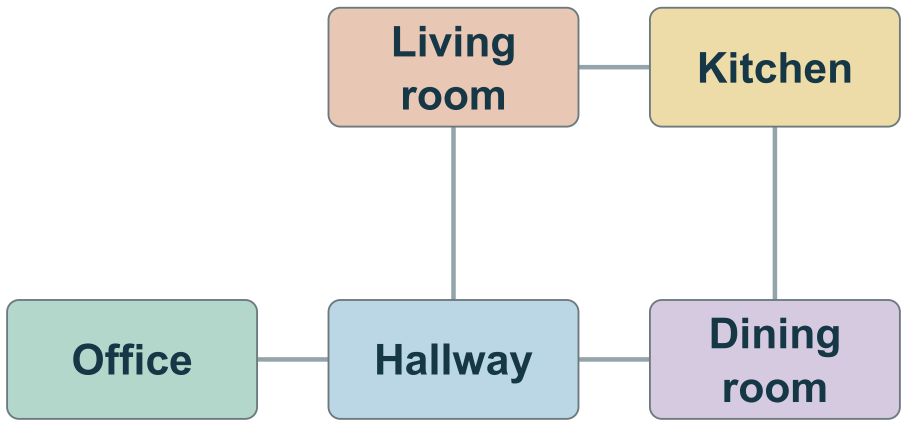
Prescribed commands. No sensor measurements yet.

Doorways—not just neighboring rooms—determine where the robot can go.



No Office ↔ Living-room doorway.

THE SAME HOUSE AS A GRAPH



Left / right / up / down are relative to this map.

Room order in every vector and matrix: L, K, O, H, D

An action advances time, even when the robot stays put.

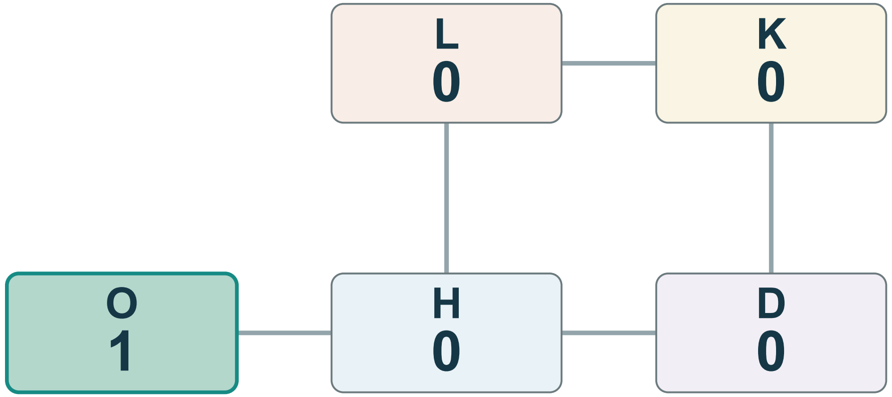
X_0		X_1		X_2
Office	a_0 = RIGHT →	**Office**	a_1 = RIGHT →	**Hallway**
start		stayed put		moved rooms

One possible trajectory—not the only outcome.

$$X_t \xrightarrow{a_t} X_{t+1}$$

The robot occupies one room; our belief assigns probabilities to all five.

KNOWN START: OFFICE



X_0 = Office

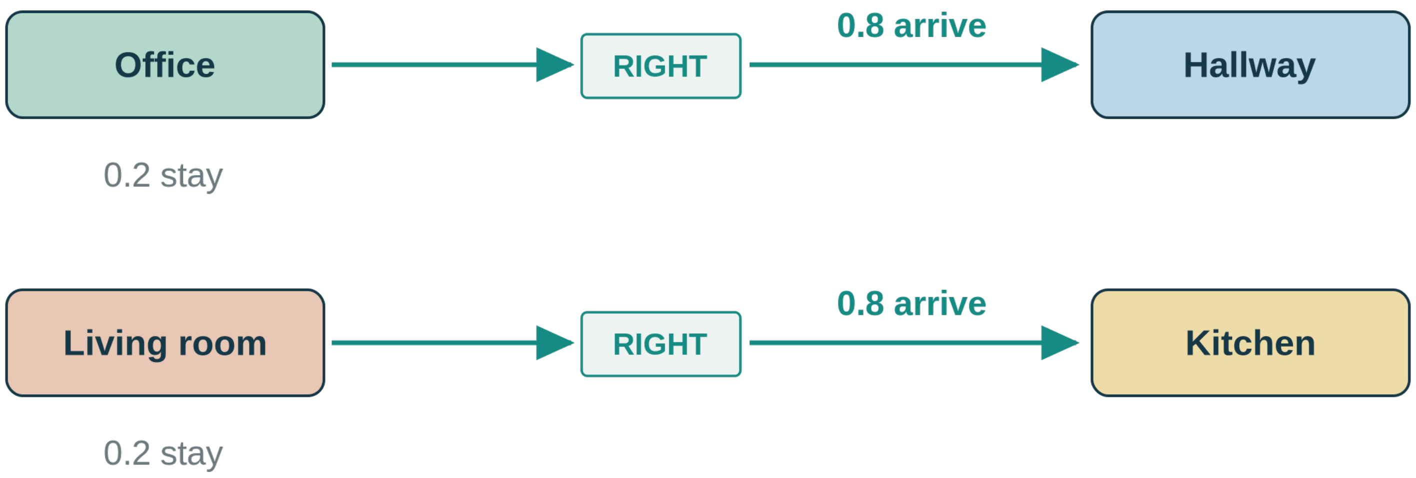
b_0: our initial belief

0	0	1	0	0
Living room	Kitchen	Office	Hallway	Dining room

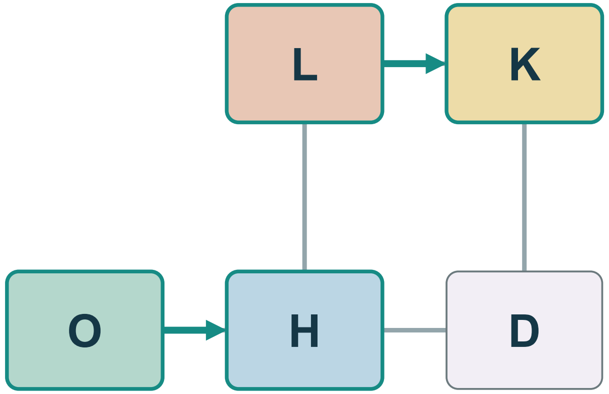
One row. Five probabilities. Total = 1.

$$b_t(x) = \text{probability assigned to room } x \text{ at step } t$$

The same command has different effects from different starting rooms.



MAP REMINDER · RIGHT



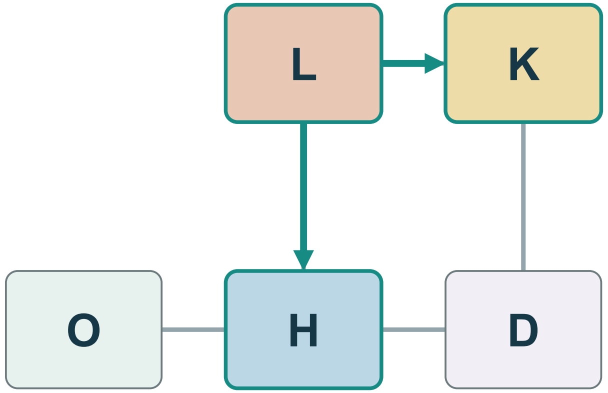
Links are doorways; arrows show the chosen direction.

Fix the current room, then choose an action; each row gives a complete distribution.

CURRENT ROOM FIXED: LIVING ROOM

ACTION	Living room	Kitchen	Office	Hallway	Dining room
LEFT	1	0	0	0	0
RIGHT	0.2	0.8	0	0	0
UP	1	0	0	0	0
DOWN	0.2	0	0	0.8	0

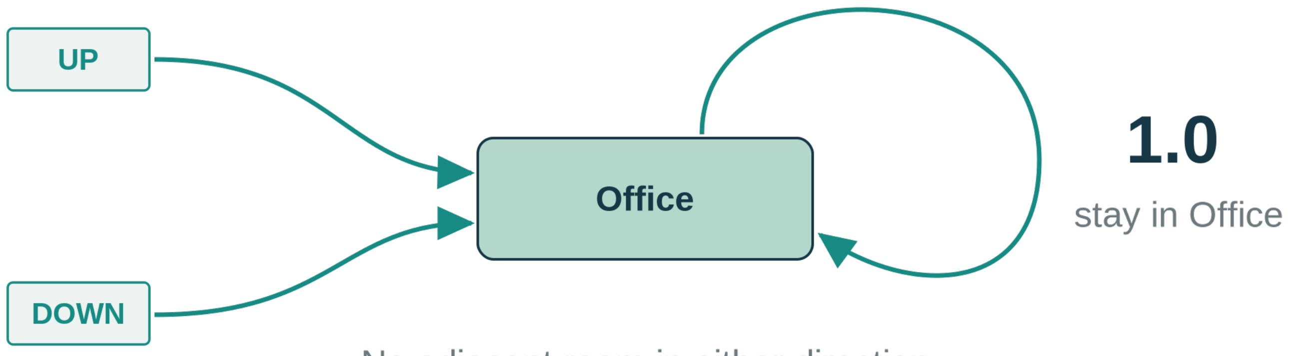
MAP REMINDER · LIVING ROOM



Links are doorways; arrows show the chosen direction.

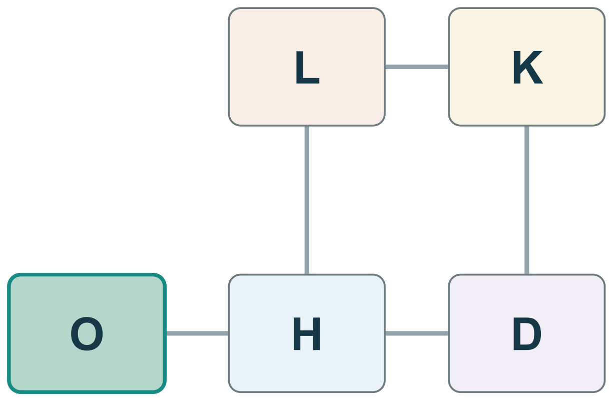
Here, rows are actions. Columns are possible next rooms. Each row totals 1.

Without a door in that direction, the robot stays in the same room.



No adjacent room in either direction.

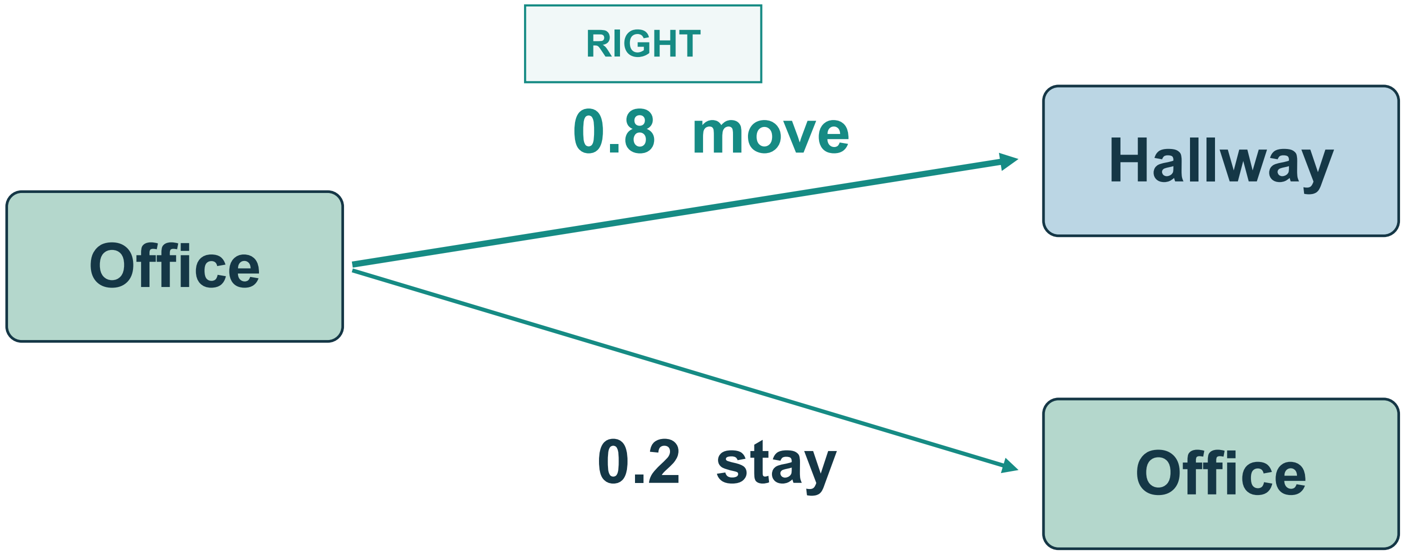
MAP REMINDER · OFFICE



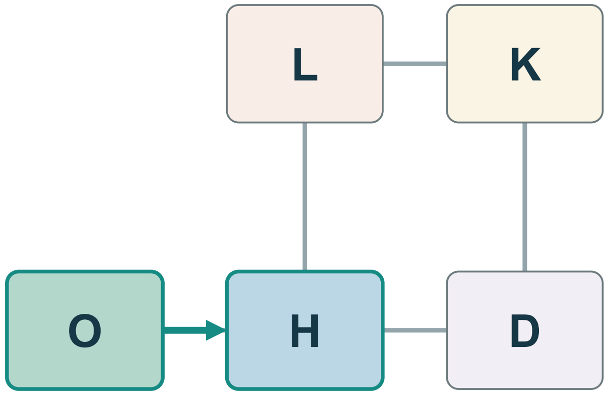
Links are doorways; arrows show the chosen direction.

No doorway: 100% stay. Available doorway: 20% stay.

One right command puts 0.8 in the Hallway and leaves 0.2 in the Office.



MAP REMINDER · RIGHT



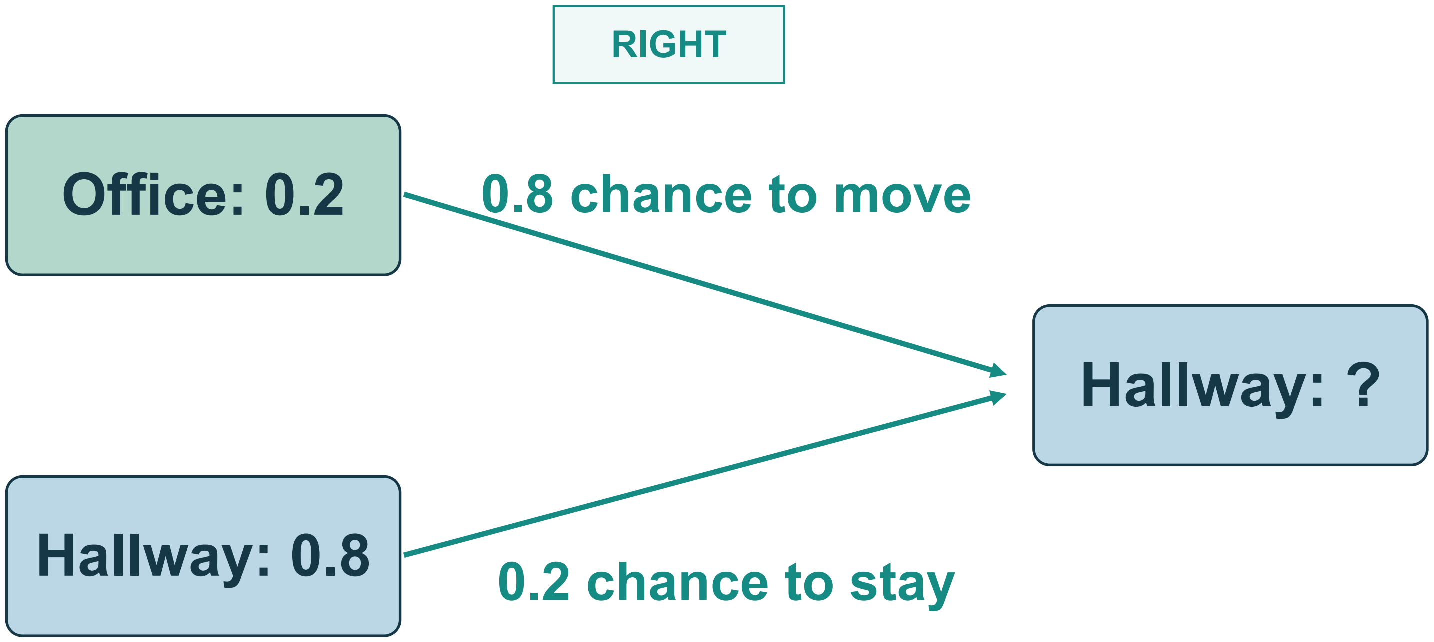
Links are doorways; arrows show the chosen direction.

$b_1 =$

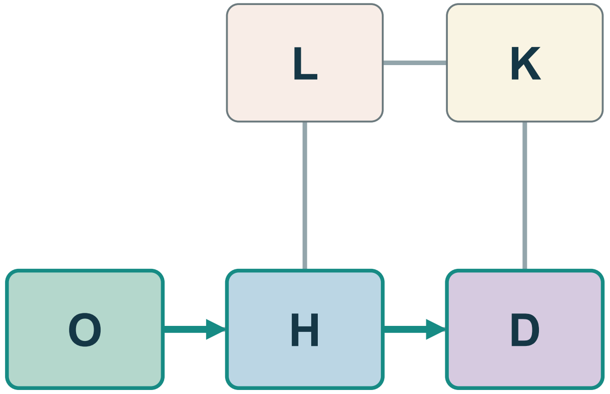
L	K	O	H	D
0	0	0.2	0.8	0

Illustrative probabilities: 80% through an available doorway; 20% stay.

Two different routes can end in the Hallway after the second command.



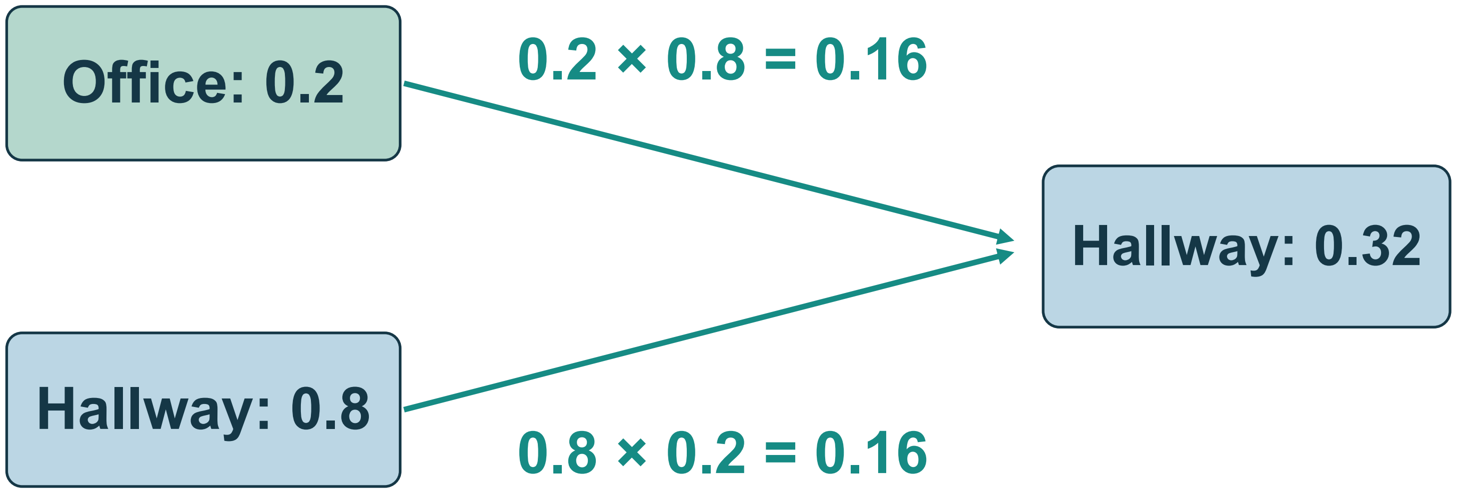
MAP REMINDER · RIGHT



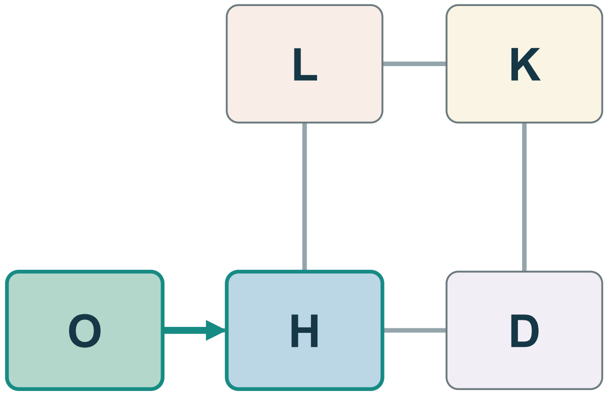
Links are doorways; arrows show the chosen direction.

What is $b_2(H)$? Work out each contribution, then combine them.

Multiply along each route; add routes that end in the same room.



MAP REMINDER · HALLWAY



Links are doorways; arrows show the chosen direction.

$$b_2(H) = 0.2 \times 0.8 + 0.8 \times 0.2 = 0.32$$

starting-room probability × conditional transition probability

For each destination, sum over every possible starting room.

$$b_{t+1}(x') = \sum_{x \in \{L,K,O,H,D\}} b_t(x)\, P(X_{t+1} = x' \mid X_t = x, a_t)$$

x = a possible starting room x′ = the destination we are updating

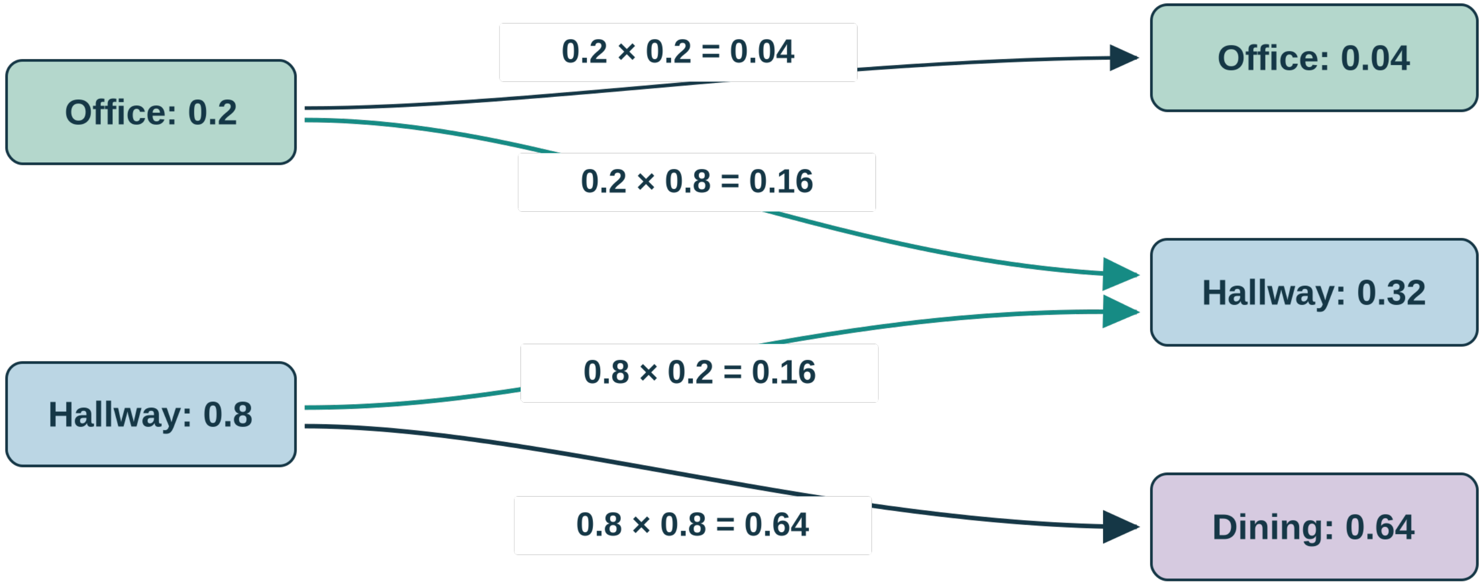
For destination H after the second RIGHT:

L		K		O		H		D
0 × 0	+	0 × 0	+	0.2 × 0.8	+	0.8 × 0.2	+	0 × 0

$$b_2(H) = 0 + 0 + 0.16 + 0.16 + 0 = 0.32$$

Some contributions are zero. We still sum over all five possible starting rooms.

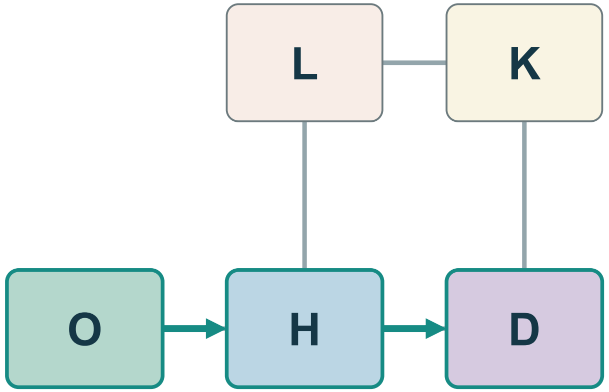
After two right commands, our belief covers three rooms—not five.



b_2 =

0	0	0.04	0.32	0.64
L	K	O	H	D

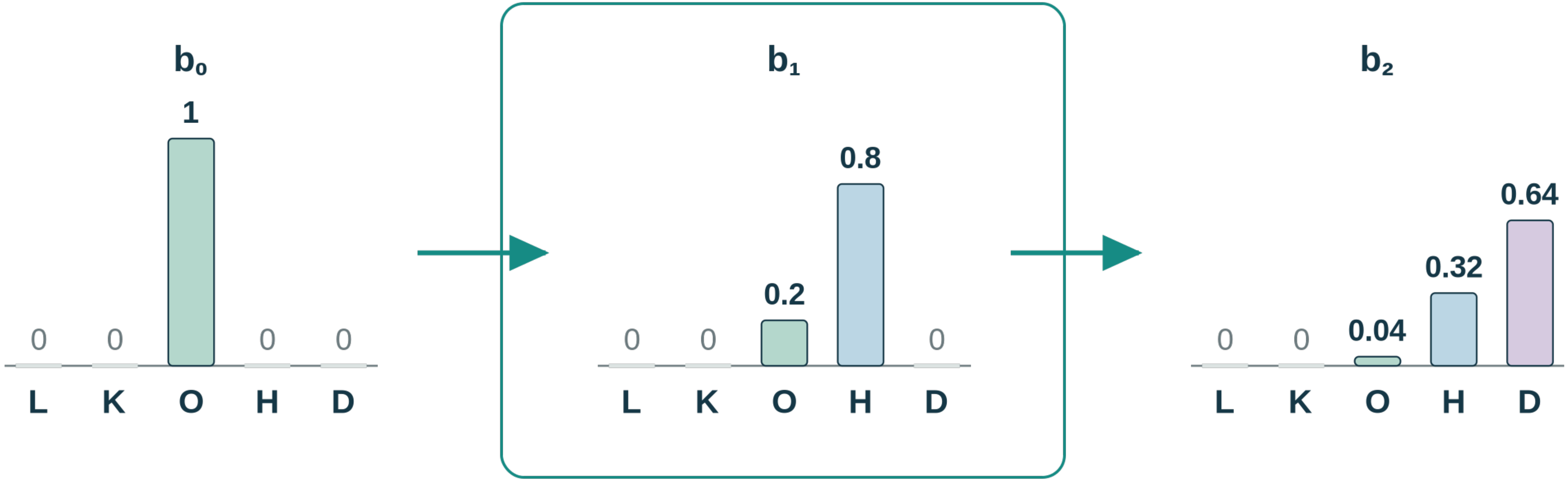
MAP REMINDER · RIGHT



Links are doorways; arrows show the chosen direction.

0.04 + 0.32 + 0.64 = 1

The result of one prediction is the starting point for the next.



posterior now = prior next

Here “posterior” means after the known actions—not after a sensor measurement.

Fix the action; collect one row from each possible starting room.

CURRENT ROOM	L	K	O	H	D	ACTION
Living room	1	0	0	0	0	L
Living room	**0.2**	**0.8**	**0**	**0**	**0**	**R**
	1	0	0	0	0	U
	0.2	0	0	0.8	0	D
Kitchen	0.8	0.2	0	0	0	L
Kitchen	**0**	**1**	**0**	**0**	**0**	**R**
	0	1	0	0	0	U
	0	0.2	0	0	0.8	D
Office	0	0	1	0	0	L
Office	**0**	**0**	**0.2**	**0.8**	**0**	**R**
	0	0	1	0	0	U
	0	0	1	0	0	D
Hallway	0	0	0.8	0.2	0	L
Hallway	**0**	**0**	**0**	**0.2**	**0.8**	**R**
	0.8	0	0	0.2	0	U
	0	0	0	1	0	D
Dining room	0	0	0	0.8	0.2	L
Dining room	**0**	**0**	**0**	**0**	**1**	**R**
	0	0.8	0	0	0.2	U
	0	0	0	0	1	D

$\longrightarrow$

M_r: move right

$$\begin{array}{c} \\ L \\ K \\ O \\ H \\ D \end{array} \begin{array}{c} \begin{array}{ccccc} L & K & O & H & D \end{array} \\ \begin{bmatrix} 0.2 & 0.8 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0.2 & 0.8 & 0 \\ 0 & 0 & 0 & 0.2 & 0.8 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \end{array}$$

One RIGHT row per starting room.

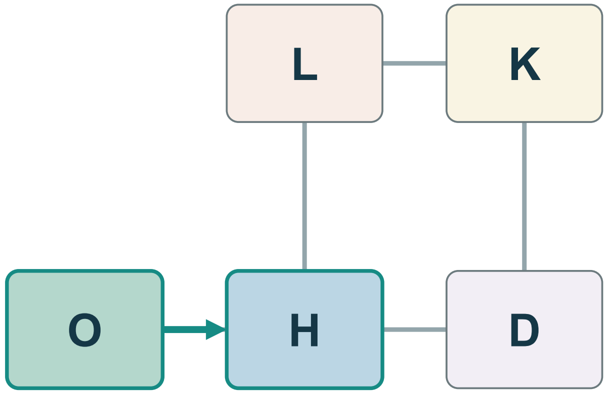
Now rows are starting rooms; columns are destination rooms.

NEXT ROOM →

CURRENT ROOM	L	K	O	H	D
Living room	**0.2**	**0.8**	0	0	0
Kitchen	0	**1**	0	0	0
Office	0	0	**0.2**	**0.8**	0
Hallway	0	0	0	**0.2**	**0.8**
Dining room	0	0	0	0	**1**

Action fixed: RIGHT

MAP REMINDER · RIGHT



Links are doorways; arrows show the chosen direction.

Row O, column H: 0.8

Each row totals one—even when the only possible outcome is staying put.

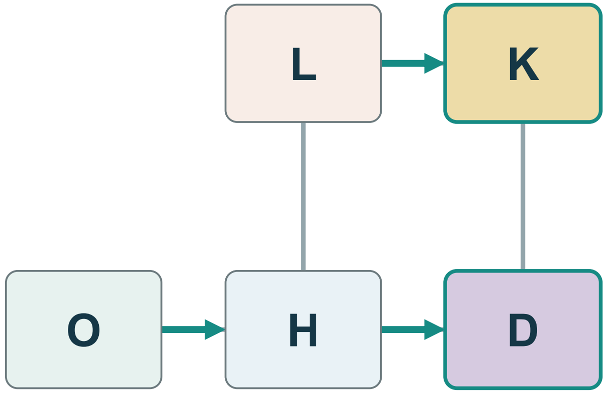
NEXT ROOM →

CURRENT ROOM	L	K	O	H	D
Living room	**0.2**	**0.8**	0	0	0
Kitchen	0	**1**	0	0	0
Office	0	0	**0.2**	**0.8**	0
Hallway	0	0	0	**0.2**	**0.8**
Dining room	0	0	0	0	**1**

Action fixed: RIGHT

Diagonal = stay in the same room.

MAP REMINDER · RIGHT



Links are doorways; arrows show the chosen direction.

K → K: 1　　D → D: 1

The Hallway column collects every way of arriving in the Hallway.

CURRENT BELIEF b_1

L	K	O	H	D
0	0	0.2	0.8	0

×

L	0
K	0
O	0.8
H	0.2
D	0

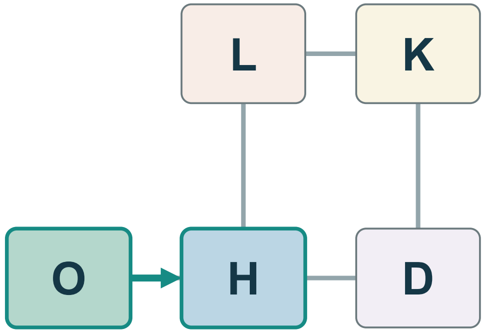
= **0.32**

$b_2(H)$

RIGHT: column H

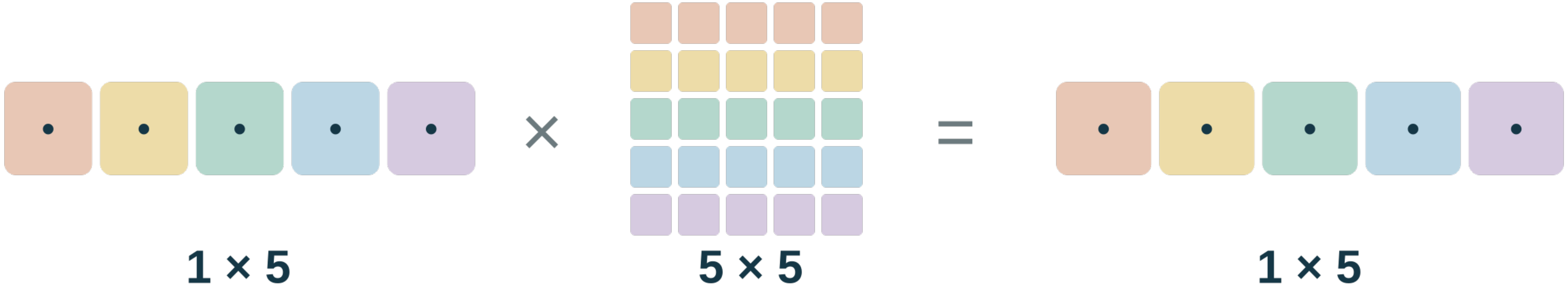
$$0 \times 0 + 0 \times 0 + 0.2 \times 0.8 + 0.8 \times 0.2 + 0 \times 0 = 0.32$$

MAP · HALLWAY



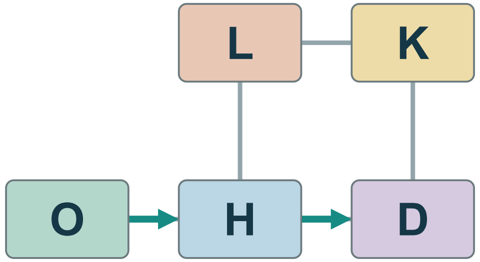
It's an elementary matrix multiply, my dear Watson.

$$b_{t+1} = b_t M_{a_t}$$



An uncertain belief mixes rows; it does not pick the most likely room.

$0.2\,\times$	0	0	0.2	0.8	0	Office row
$+\ 0.8\,\times$	0	0	0	0.2	0.8	Hallway row
$b_2 =$	0	0	0.04	0.32	0.64	
	L	K	O	H	D	



Keep every current-room hypothesis, weighted by its probability.

What happens to probability of the Dining room keeps when we command right again?

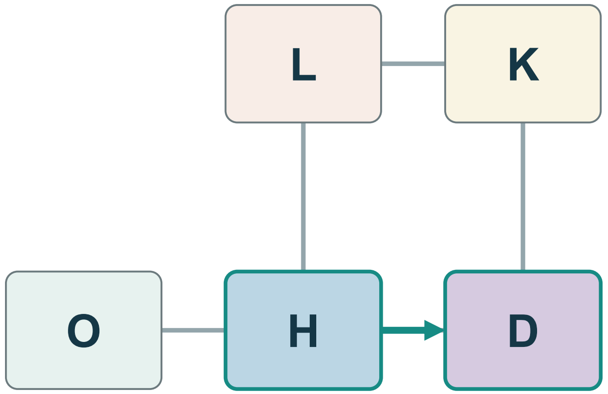
L	K	O	H	D
0	0	0.04	0.32	0.64

Current belief b_2

Hallway: 0.32 $\xrightarrow{0.8}$ **Dining: 0.64**

stays with probability 1

MAP REMINDER · RIGHT

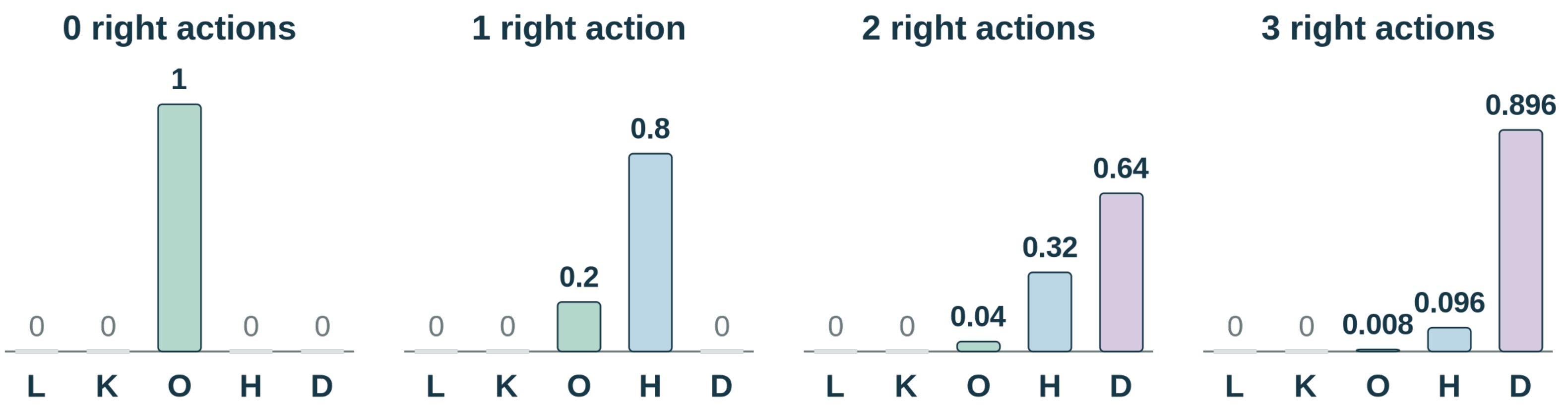


Links are doorways; arrows show the chosen direction.

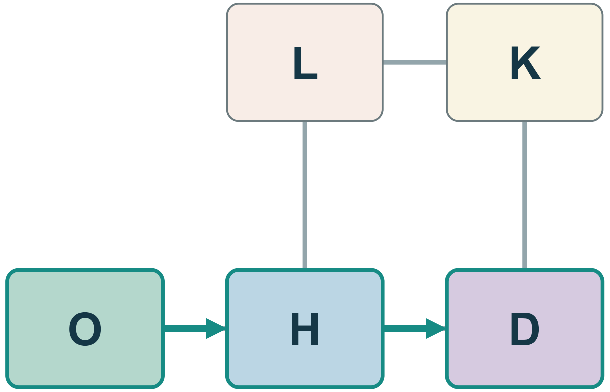
What is $b_3(D)$?

Repeating the same action means taking a matrix power.

$$b_n = b_0 M_R^n$$



MAP REMINDER · RIGHT



Links are doorways; arrows show the chosen direction.

$b_3(D) = 0.64 \times 1 + 0.32 \times 0.8 = 0.896$

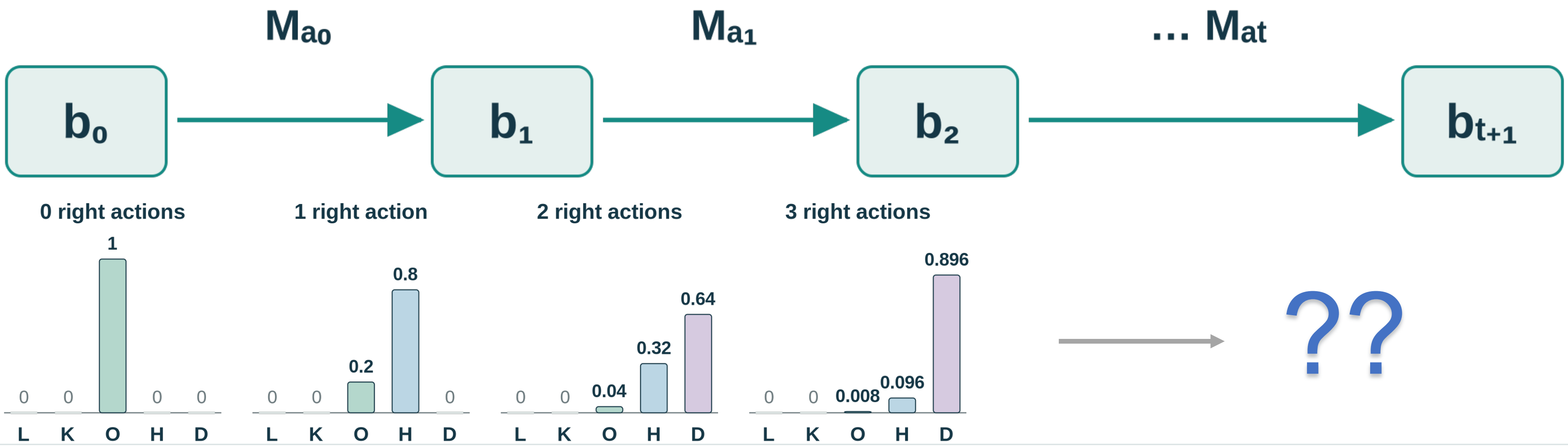
A sequence of actions becomes a sequence of matrix multiplies.

$$b_{t+1} = b_0 M_{a_0} M_{a_1} \cdots M_{a_t}$$

Apply the matrices in the order of the actions.

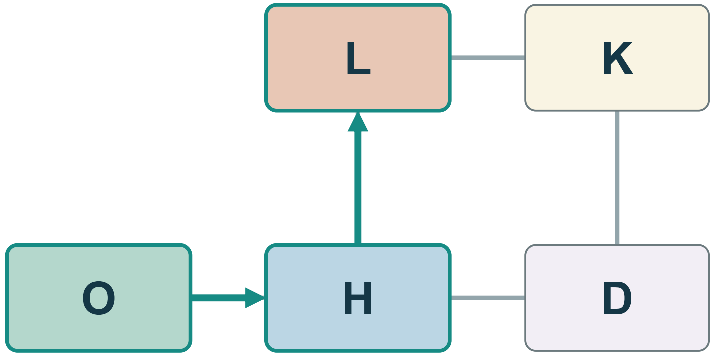
Aside: Google Page-Rank applies same M ad infinitum

$$b_{t+1} = b_0 M_{a_0} M_{a_1} \cdots M_{a_t}$$



Changing the action order changes where the robot can end up.

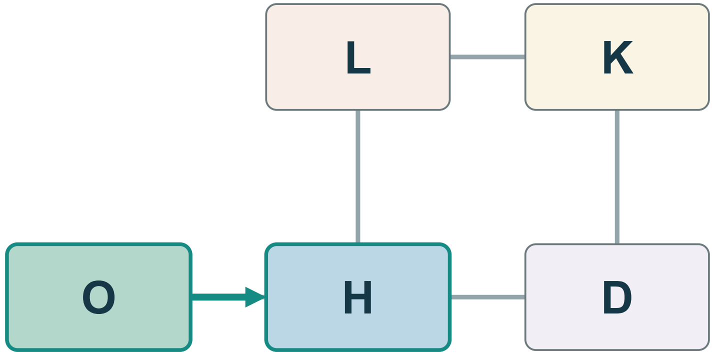
RIGHT → UP



1. O → H   2. H → L

L	K	O	H	D
0.64	0	0.2	0.16	0

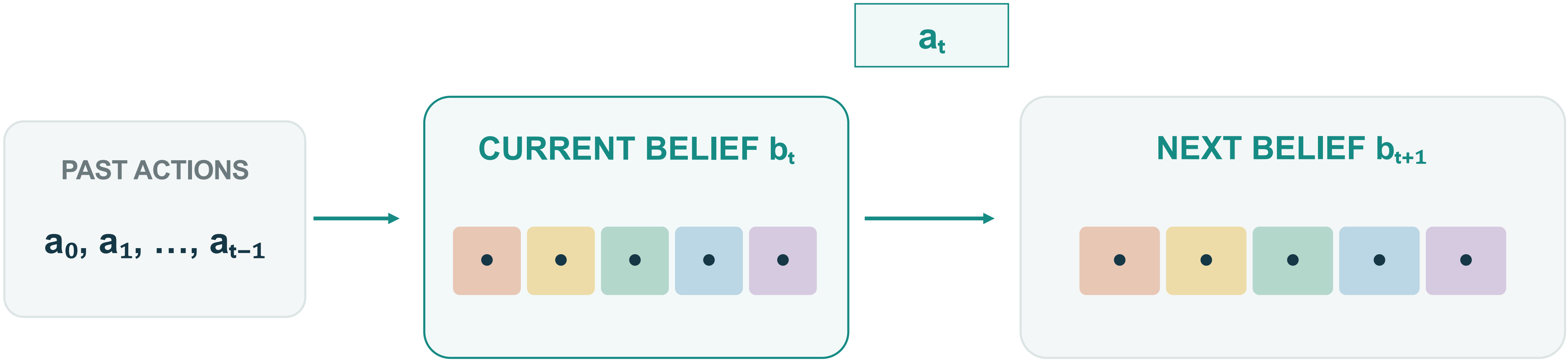
UP → RIGHT



1. UP stays in O   2. O → H

L	K	O	H	D
0	0	0.2	0.8	0

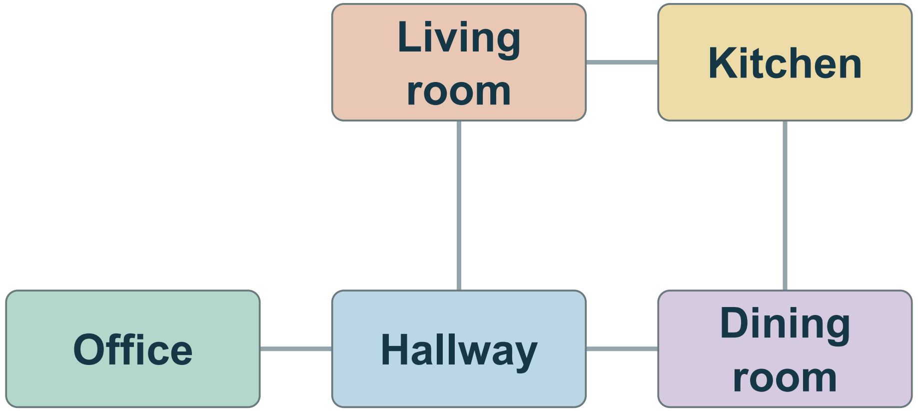
The current belief summarizes the history needed for prediction.



$$b_{t+1} = b_t M_{a_t}$$

A room graph and a time graph answer different questions.

WHERE CAN THE ROBOT GO?



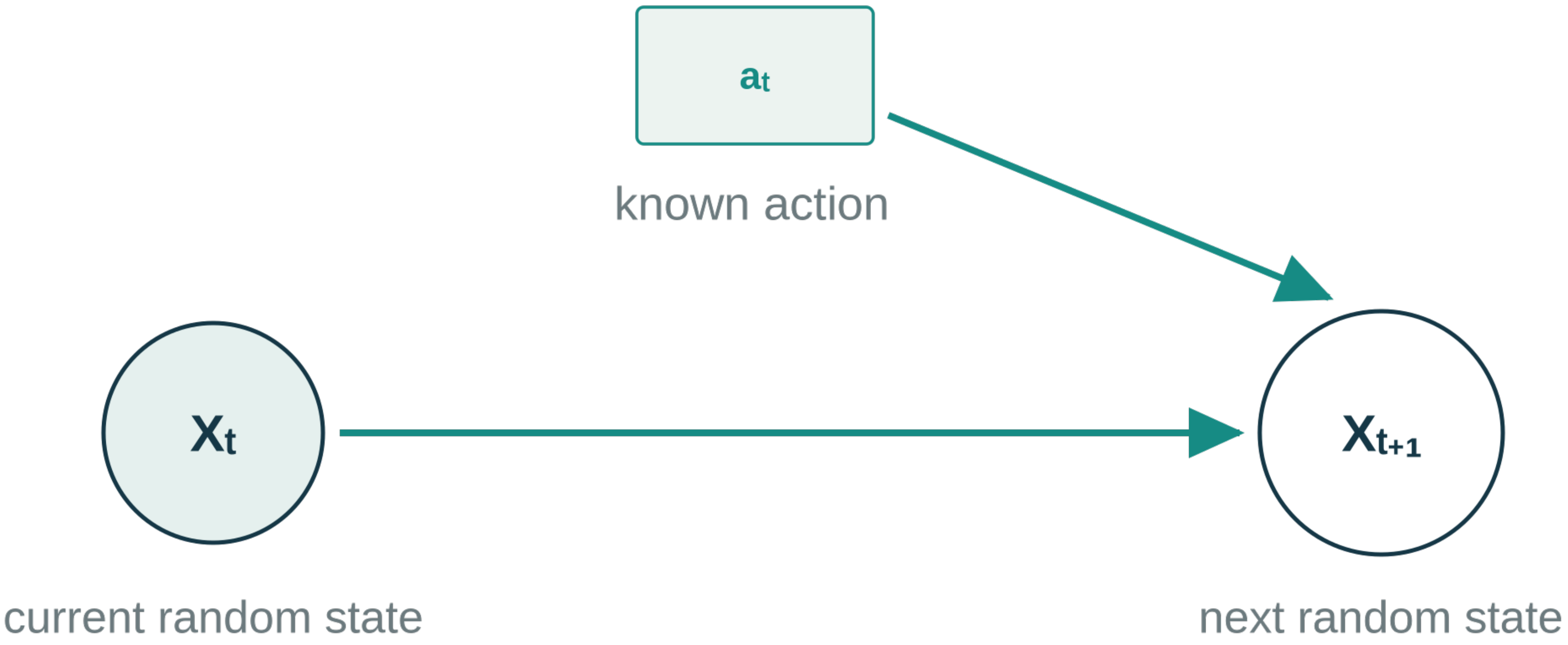
Node = a room
Link = a doorway

HOW DOES THE STATE EVOLVE?

Each X_t can take any of the five room values.

Node = the state at a time
Arrow = conditional dependence

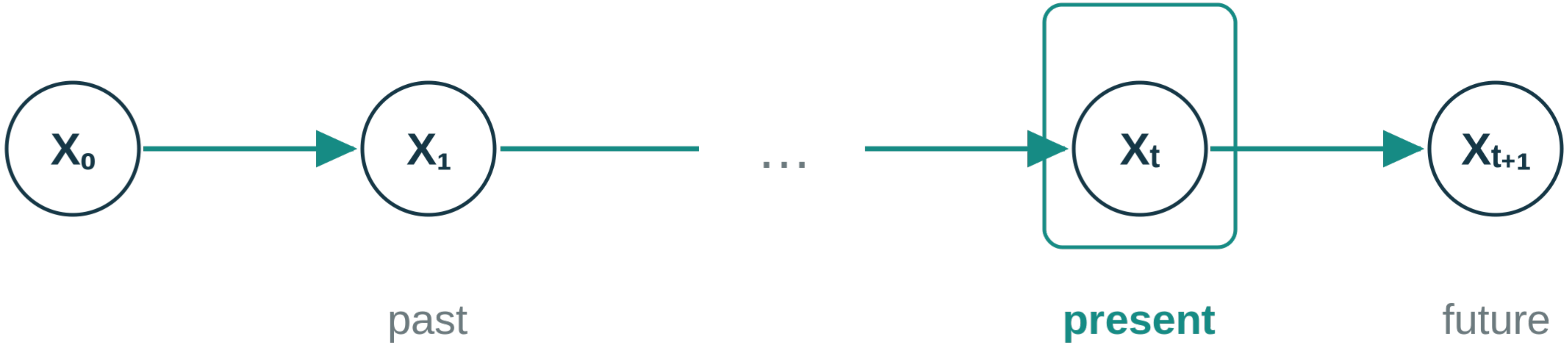
The next state depends on the current state and the chosen action.



$$P(X_{t+1} \mid X_t, a_t)$$

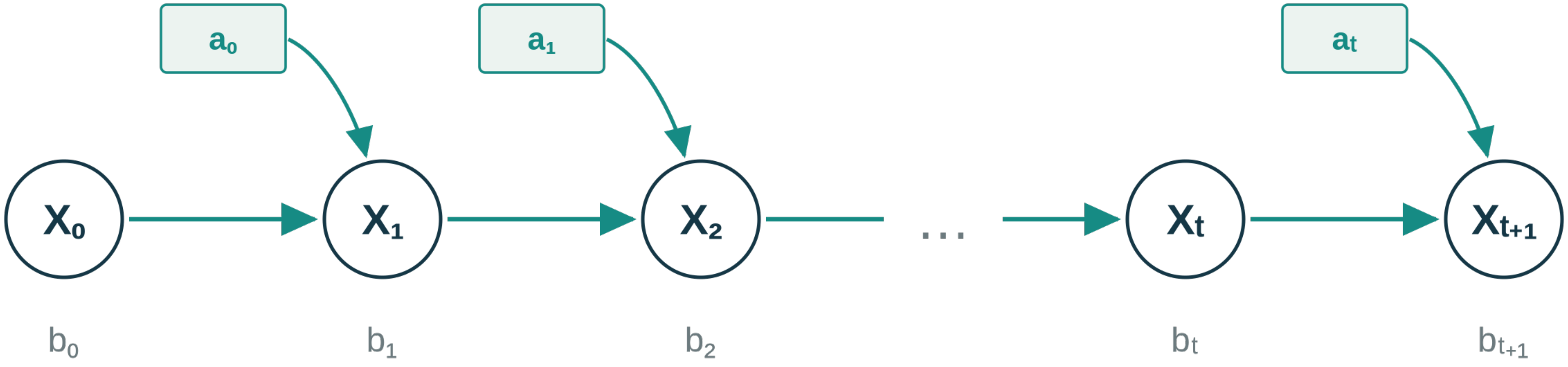
With the actions fixed, the state sequence forms a Markov chain.

$$P(X_{t+1} \mid X_0, \ldots, X_t) = P(X_{t+1} \mid X_t)$$



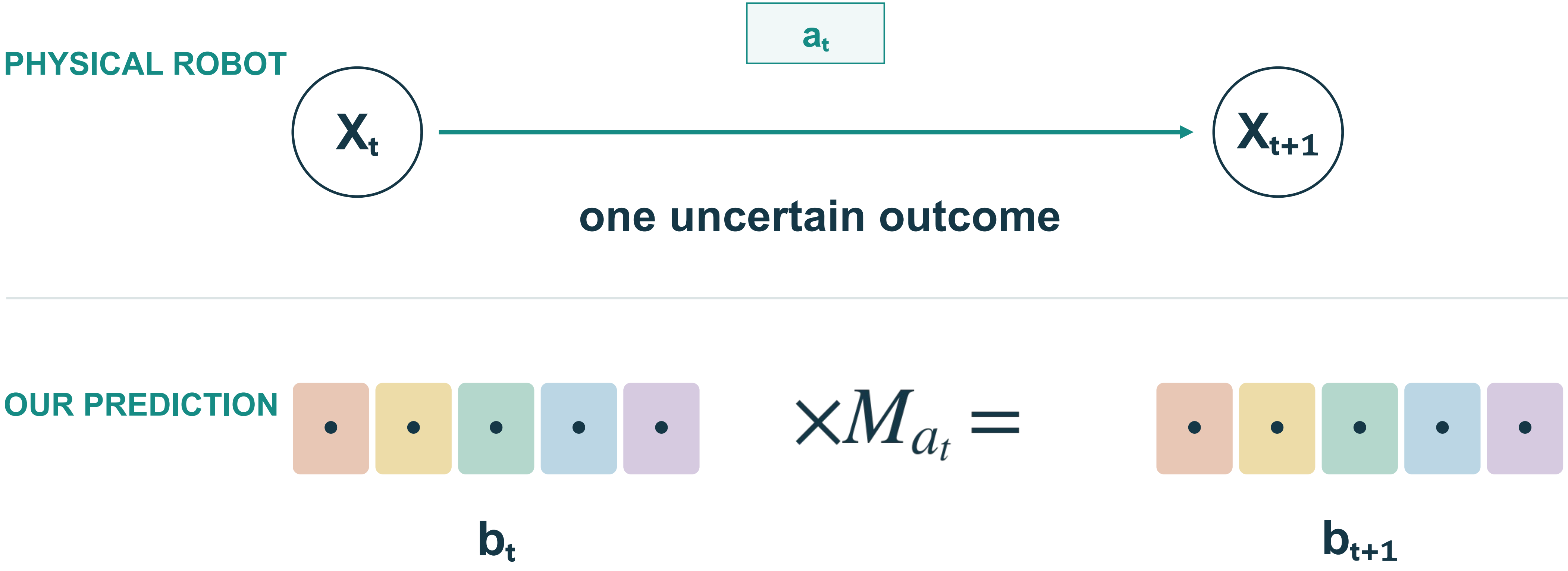
For a fixed action sequence, the robot’s states form a Markov chain.

Markov chains are everywhere.
Actions control them.



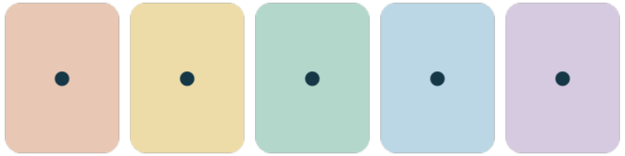
Each action selects the conditional distribution for the next transition.

The next room is random; the next belief is a calculation.



Choose the action → choose the matrix, not the realized next room.

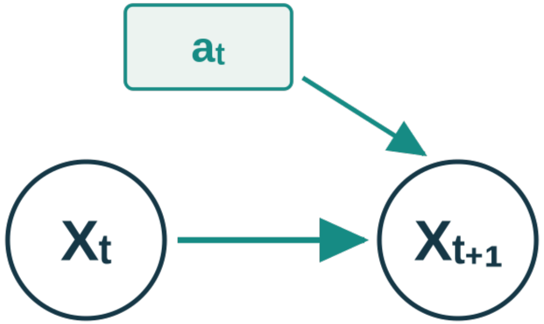
An action model and a belief state are enough to predict what comes next.



KEEP A BELIEF

$$\begin{bmatrix} 0.2 & 0.8 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0.2 & 0.8 & 0 \\ 0 & 0 & 0 & 0.2 & 0.8 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

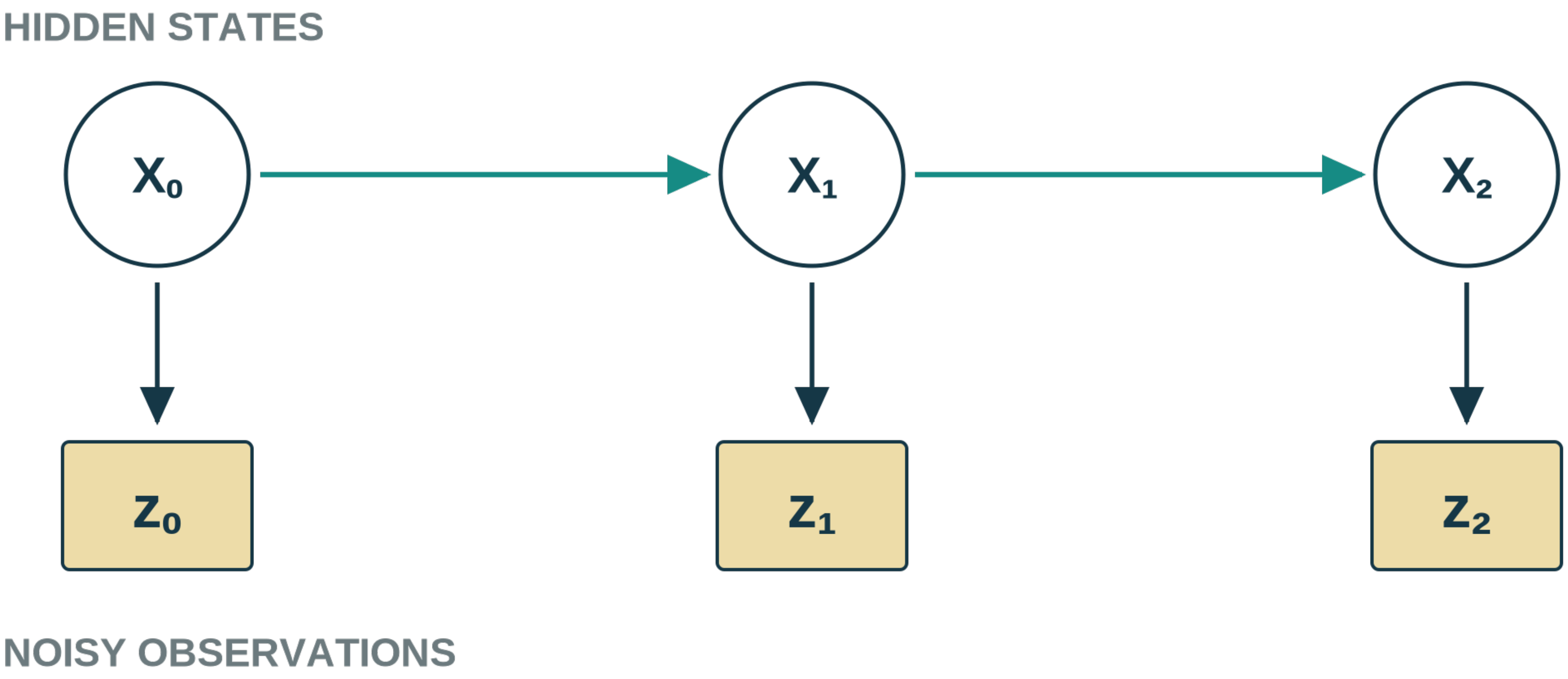
APPLY AN ACTION MODEL



PREDICT THE NEXT STATE

No need to replay the entire past.

Next, noisy measurements will help us infer the hidden state.



Bayes nets → sensing over time → hidden Markov models.

The belief vector is shorthand for a distribution conditioned on past actions.

$b_t(L)$	$b_t(K)$	$b_t(O)$	$b_t(H)$	$b_t(D)$
Living room	**Kitchen**	**Office**	**Hallway**	**Dining room**

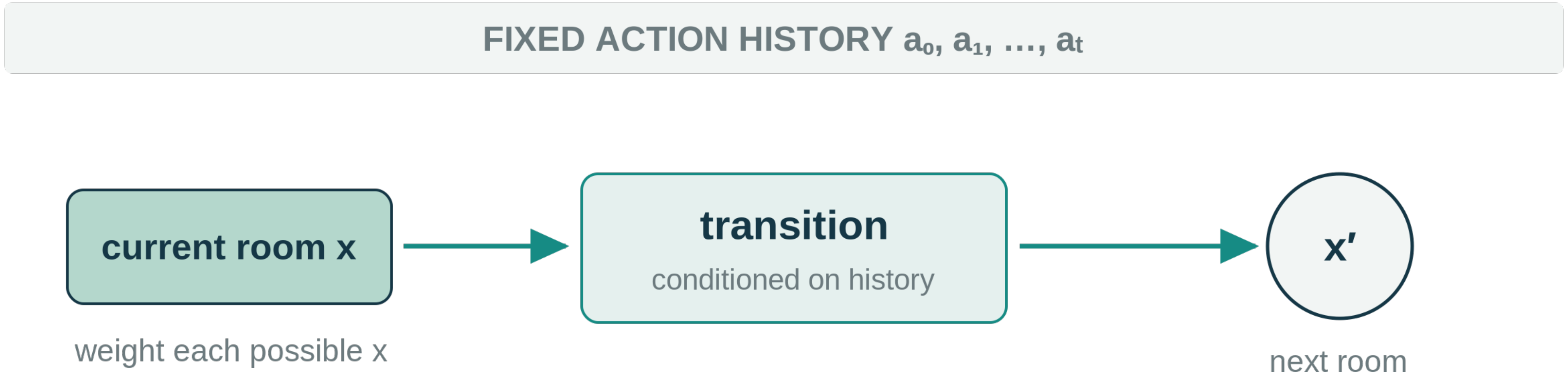
One row. Five probabilities. Total mass = 1.

$$b_t(x) = P(X_t = x \mid a_{0:t-1}) \quad (t \geq 1)$$

$$b_0(x) = P(X_0 = x)$$

Condition on the action history, then sum over the current room.

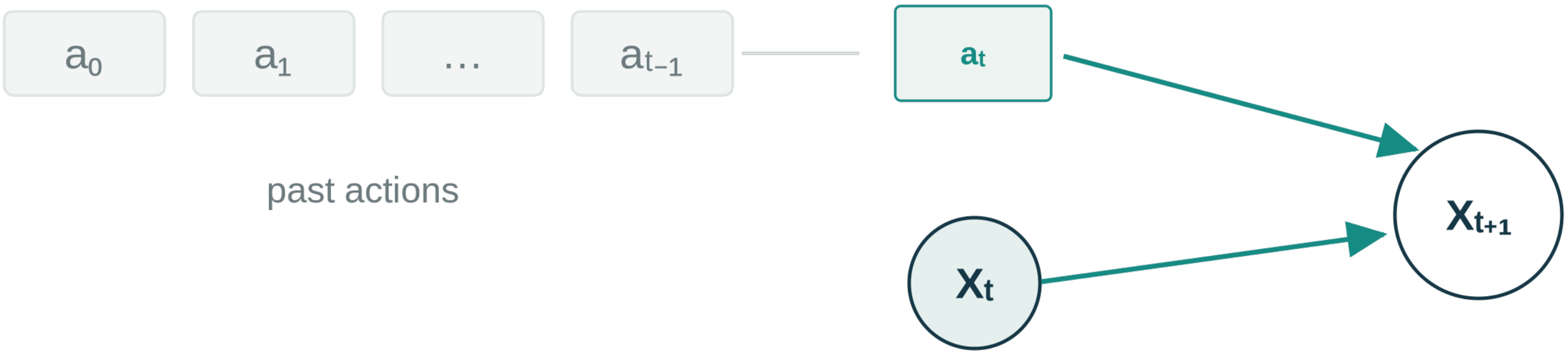
$$P(X_{t+1} = x' \mid a_{0:t}) = \sum_x P(X_{t+1} = x' \mid X_t = x, a_{0:t}) \cdot P(X_t = x \mid a_{0:t})$$



$a_{0:t}$ abbreviates the prescribed actions $a_0, \ldots, a_t$.

Given the current room and action, earlier actions drop out of the transition.

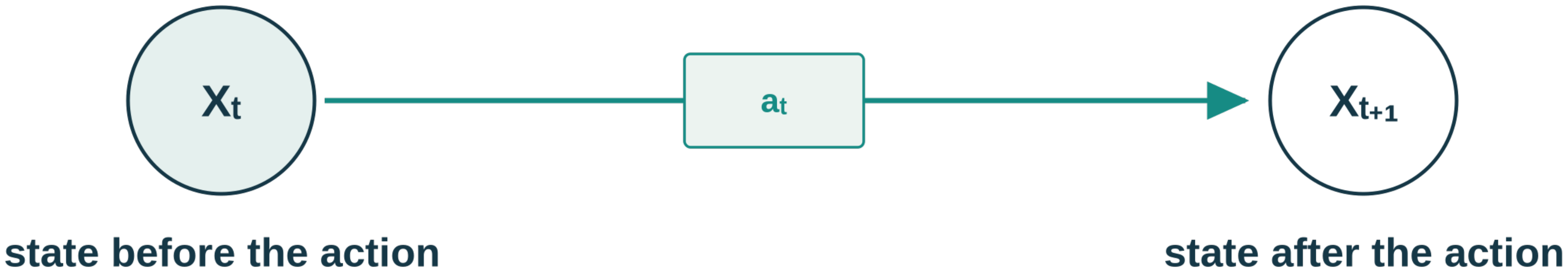
$$P(X_{t+1} = x' \mid X_t = x, a_{0:t}) = P(X_{t+1} = x' \mid X_t = x, a_t)$$



Only the present is needed.

A command chosen in advance adds no information about the starting room.

$$P(X_t = x \mid a_{0:t}) = P(X_t = x \mid a_{0:t-1})$$

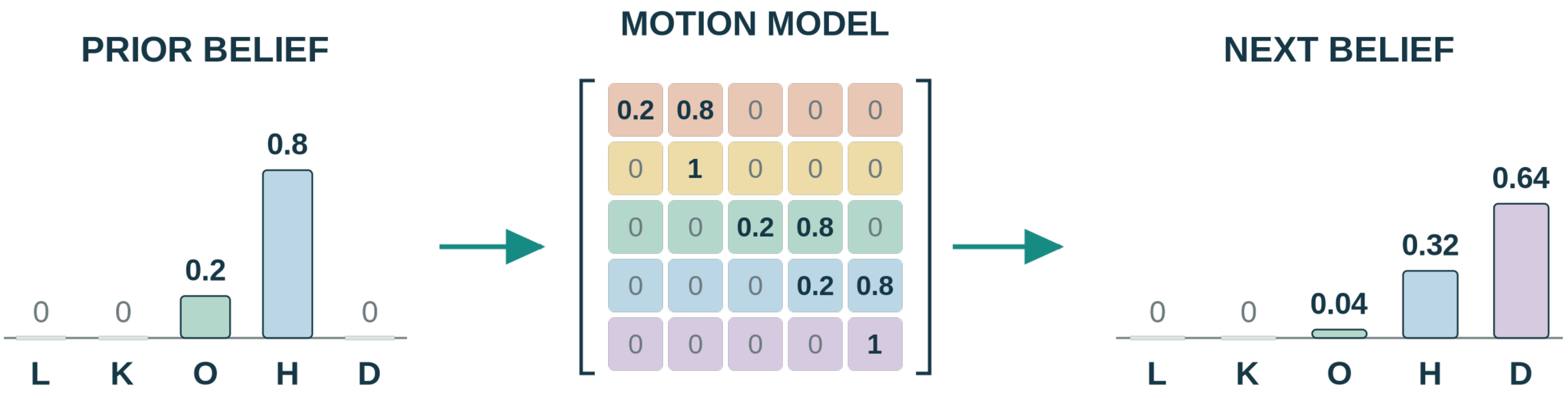


The action is applied between these states.

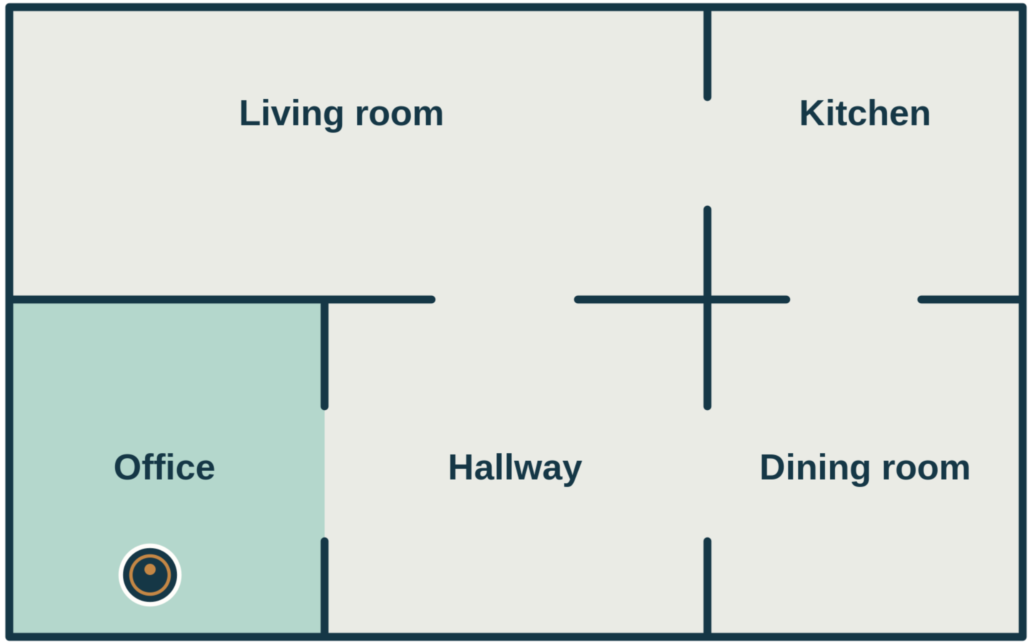
For the fixed action sequence considered in this lecture.

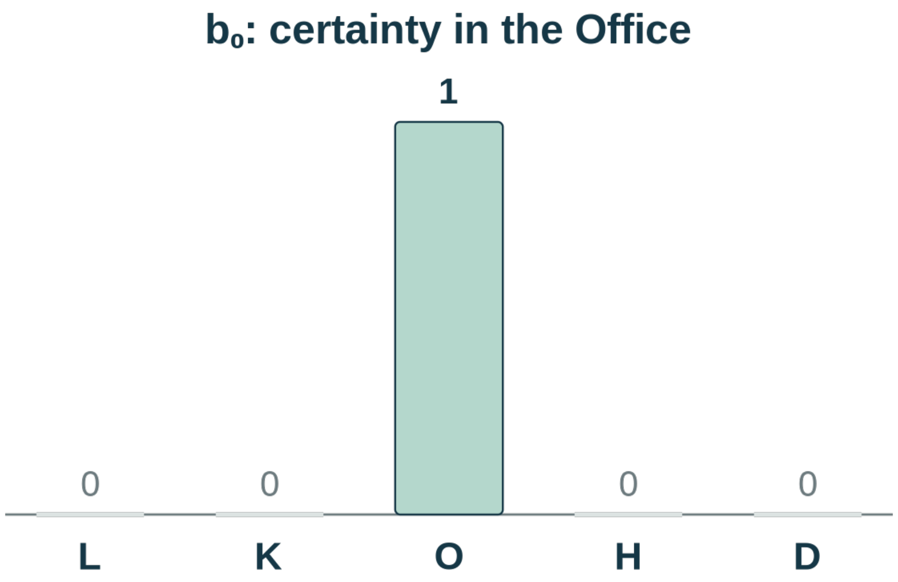
Substituting the two simplifications gives the same belief update.

$$P(X_{t+1} = x' \mid a_{0:t}) = \sum_{x} P(X_{t+1} = x' \mid X_t = x, a_t)\, P(X_t = x \mid a_{0:t-1})$$



The first update starts with the initial distribution.





$$P(X_1 = x' \mid a_0) = \sum_x P(X_1 = x' \mid X_0 = x, a_0)\, P(X_0 = x)$$

The full motion model fits into four transition matrices.

LEFT

$$\begin{array}{c|c} & \begin{matrix} L & K & O & H & D \end{matrix} \\ \begin{matrix} L \\ K \\ O \\ H \\ D \end{matrix} & \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0.8 & 0.2 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0.8 & 0.2 & 0 \\ 0 & 0 & 0 & 0.8 & 0.2 \end{bmatrix} \end{array}$$

RIGHT

$$\begin{array}{c|c} & \begin{matrix} L & K & O & H & D \end{matrix} \\ \begin{matrix} L \\ K \\ O \\ H \\ D \end{matrix} & \begin{bmatrix} 0.2 & 0.8 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0.2 & 0.8 & 0 \\ 0 & 0 & 0 & 0.2 & 0.8 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \end{array}$$

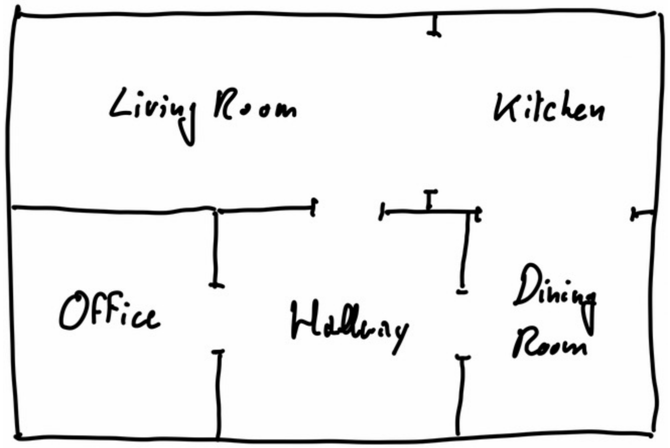
UP

$$\begin{array}{c|c} & \begin{matrix} L & K & O & H & D \end{matrix} \\ \begin{matrix} L \\ K \\ O \\ H \\ D \end{matrix} & \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0.8 & 0 & 0 & 0.2 & 0 \\ 0 & 0.8 & 0 & 0 & 0.2 \end{bmatrix} \end{array}$$

DOWN

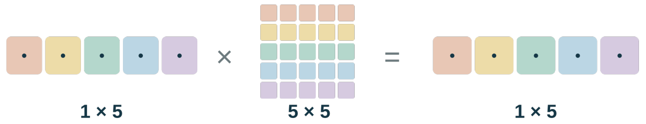
$$\begin{array}{c|c} & \begin{matrix} L & K & O & H & D \end{matrix} \\ \begin{matrix} L \\ K \\ O \\ H \\ D \end{matrix} & \begin{bmatrix} 0.2 & 0 & 0 & 0.8 & 0 \\ 0 & 0.2 & 0 & 0 & 0.8 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \end{array}$$

Section 3.2 connects action uncertainty to controlled Markov chains.



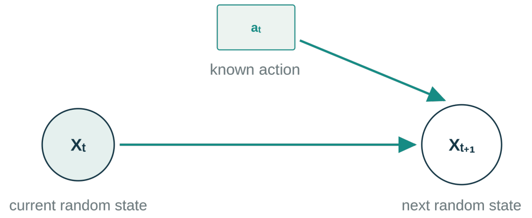
§3.2.1

Probabilistic action outcomes



§§3.2.2–3.2.3

Belief updates and transition matrices



§3.2.4

Controlled Markov chains