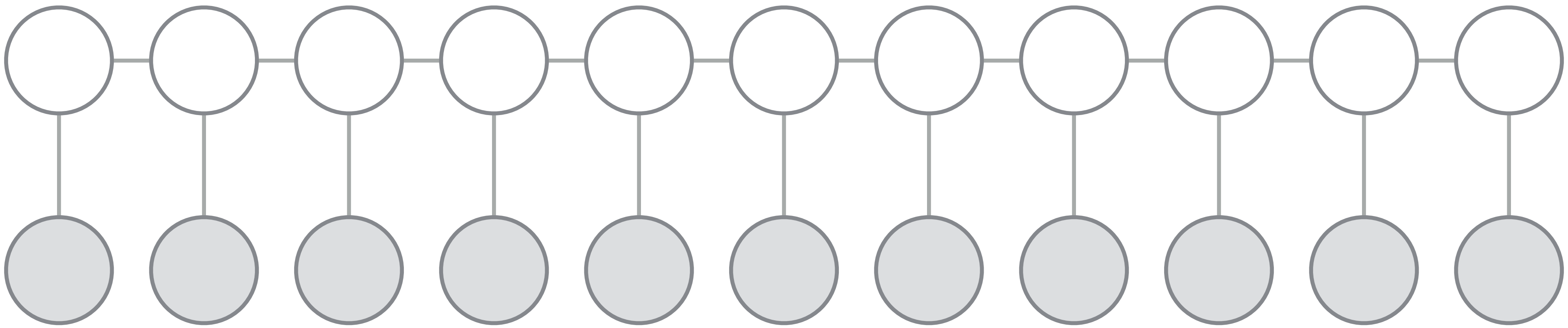


Temporal State Models

16-385 Computer Vision (Kris Kitani)

Carnegie Mellon University

Represent the 'world' as a set of random variables $\boldsymbol{X}$

$\boldsymbol{X} = \{\boldsymbol{x}, \boldsymbol{y}\}$ location on the ground plane

$\boldsymbol{X} = \{\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}\}$ position in the 3D world

$\boldsymbol{X} = \{\boldsymbol{x}, \dot{\boldsymbol{x}}\}$ position and velocity

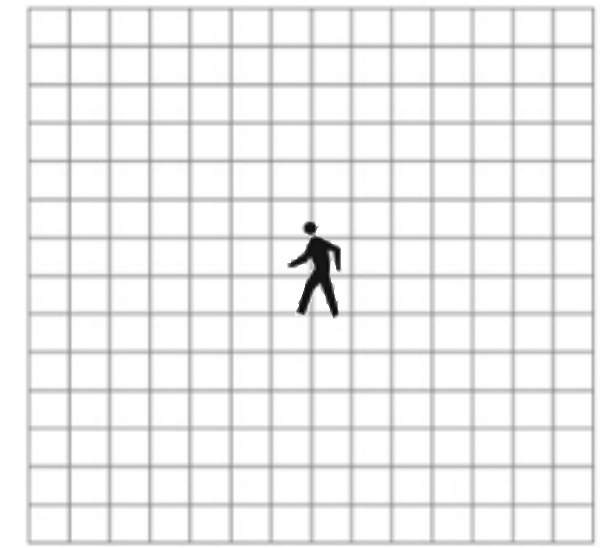
$\boldsymbol{X} = \{\boldsymbol{x}, \dot{\boldsymbol{x}}, \boldsymbol{f}_1, \ldots, \boldsymbol{f}_n\}$

position, velocity and
location of landmarks

Object tracking (localization)

$$\boldsymbol{X} = \{\boldsymbol{x}, \boldsymbol{y}\}$$

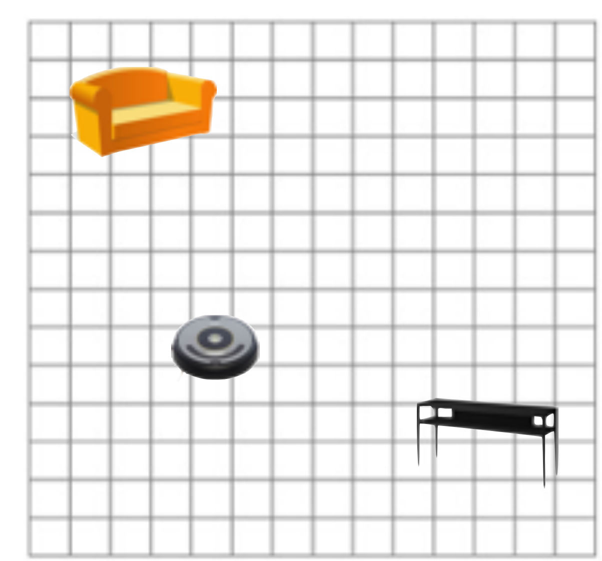
e.g., location on the ground plane



Object location and world landmarks (localization and mapping)

$$\boldsymbol{X} = \{\boldsymbol{x}, \dot{\boldsymbol{x}}, \boldsymbol{f}_1, \ldots, \boldsymbol{f}_n\}$$

e.g., position and velocity of robot and location of landmarks



$$\boldsymbol{X}_t$$

The state of the world changes over time

$$\boldsymbol{X}_t$$

The state of the world changes over time

So we use a sequence of random variables:

$$\boldsymbol{X}_0, \boldsymbol{X}_1, \ldots, \boldsymbol{X}_t$$

$$\boldsymbol{X}_t$$

The state of the world changes over time

So we use a sequence of random variables:

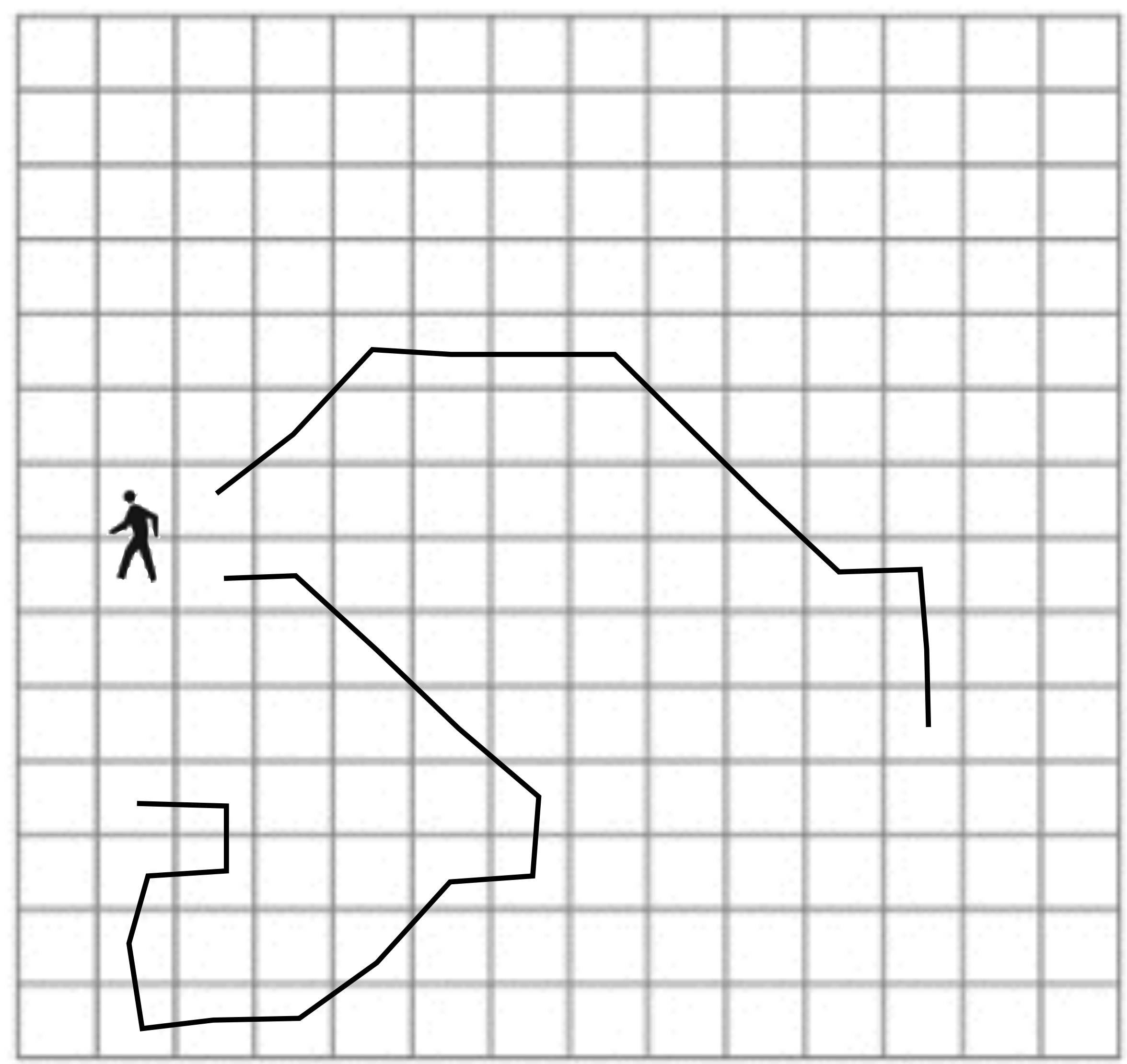
$$\boldsymbol{X}_0, \boldsymbol{X}_1, \ldots, \boldsymbol{X}_t$$

The state of the world is usually **uncertain** so we think in terms of a distribution

$$P(\boldsymbol{X}_0, \boldsymbol{X}_1, \ldots, \boldsymbol{X}_t)$$

How big is the space of this distribution?

If the state space is $\boldsymbol{X} = \{\boldsymbol{x}, \boldsymbol{y}\}$ the location on the ground plane



$$P(\boldsymbol{X}_0, \boldsymbol{X}_1, \ldots, \boldsymbol{X}_t)$$

is the probability over all possible trajectories through a room of length t+1

When we use a sensor (camera),
we don't have direct access to the state but noisy observations of the state

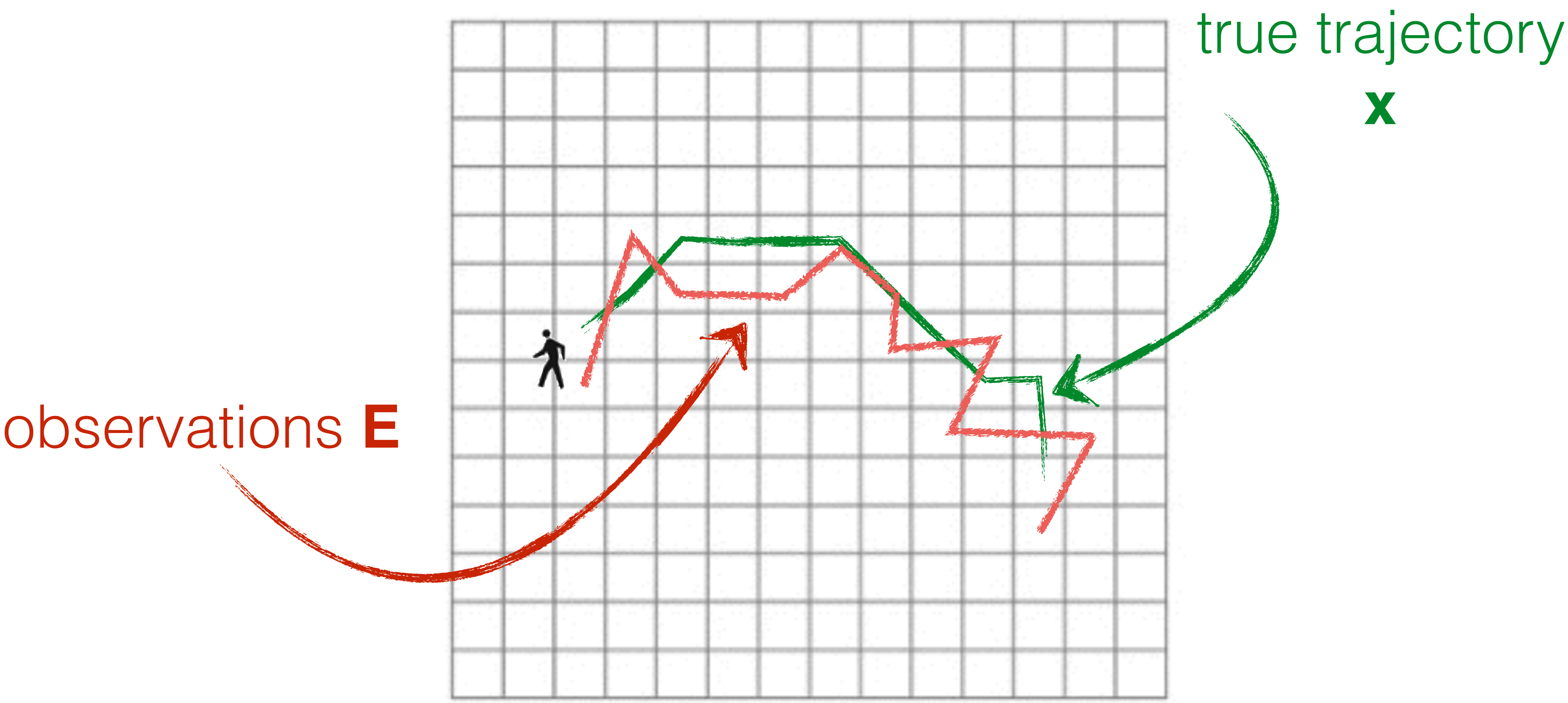
$$\boldsymbol{E}_t$$

$$\boldsymbol{X}_0, \boldsymbol{X}_1, \ldots, \boldsymbol{X}_t, \boldsymbol{E}_1, \boldsymbol{E}_2, \ldots, \boldsymbol{E}_t$$

(**all** possible ways of observing **all** possible trajectories)

How big is the space of this distribution?

all possible ways of observing **all** possible trajectories of length t



So we think of the world in terms of the distribution

$$P(\underbrace{\boldsymbol{X}_0, \boldsymbol{X}_1, \ldots, \boldsymbol{X}_t}_{\text{unobserved variables (hidden state)}}, \underbrace{\boldsymbol{E}_1, \boldsymbol{E}_2, \ldots, \boldsymbol{E}_t}_{\text{observed variables (evidence)}})$$

So we think of the world in terms of the distribution

$$P(\underbrace{\boldsymbol{X}_0, \boldsymbol{X}_1, \ldots, \boldsymbol{X}_t}_{\substack{\text{unobserved variables} \\ \text{(hidden state)}}}, \underbrace{\boldsymbol{E}_1, \boldsymbol{E}_2, \ldots, \boldsymbol{E}_t}_{\substack{\text{observed variables} \\ \text{(evidence)}}})$$

How big is the space of this distribution?

So we think of the world in terms of the distribution

$$P(\underbrace{\boldsymbol{X}_0, \boldsymbol{X}_1, \ldots, \boldsymbol{X}_t}_{\text{unobserved variables (hidden state)}}, \underbrace{\boldsymbol{E}_1, \boldsymbol{E}_2, \ldots, \boldsymbol{E}_t}_{\text{observed variables (evidence)}})$$

How big is the space of this distribution?

Can you think of a way to reduce the space?

Reduction 1. Stationary process assumption:

‘a process of change that is governed by laws that do not themselves change over time.’

$$P(\boldsymbol{E}_t|\boldsymbol{X}_t) = P_t(\boldsymbol{E}_t|\boldsymbol{X}_t)$$

the model doesn’t change over time

Reduction 1. Stationary process assumption:

'a process of change that is governed by laws that do not themselves change over time.'

$$P(\boldsymbol{E}_t|\boldsymbol{X}_t) = P_t(\boldsymbol{E}_t|\boldsymbol{X}_t)$$

the model doesn't change over time

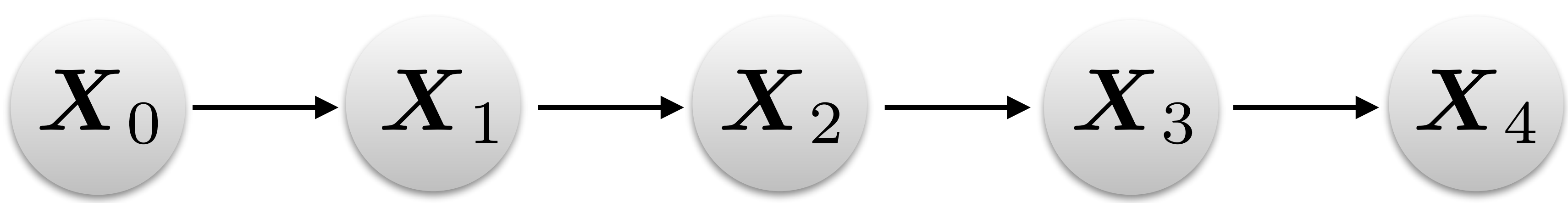
Only have to store **one** model.

Is this a reasonable assumption?

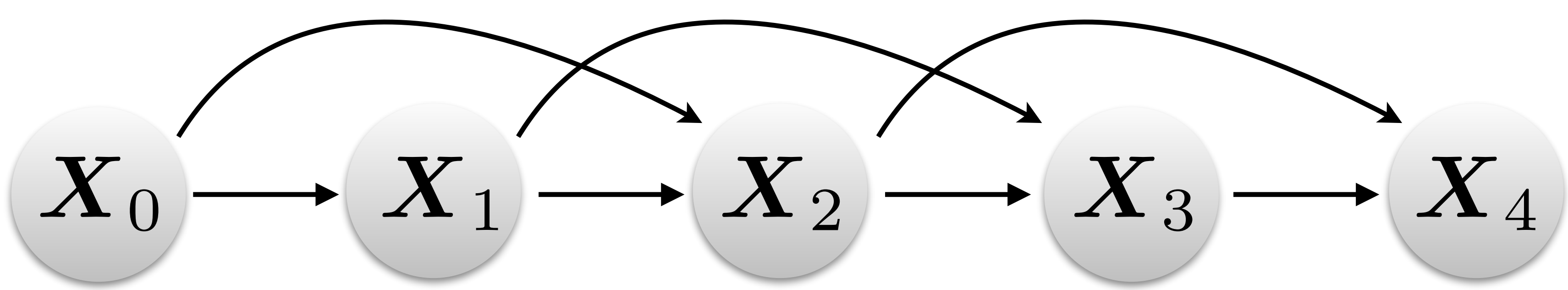
Reduction 2. Markov Assumption:

'the current state only depends on a finite history of previous states.'

First-order Markov Model: $P(\boldsymbol{X}_t|\boldsymbol{X}_{t-1})$.



Second-order Markov Model: $P(\boldsymbol{X}_t|\boldsymbol{X}_{t-1}, \boldsymbol{X}_{t-2})$



(this relationship is called the **motion** model)

Reduction 2. Markov Assumption:

'the current observation only depends on current state.'

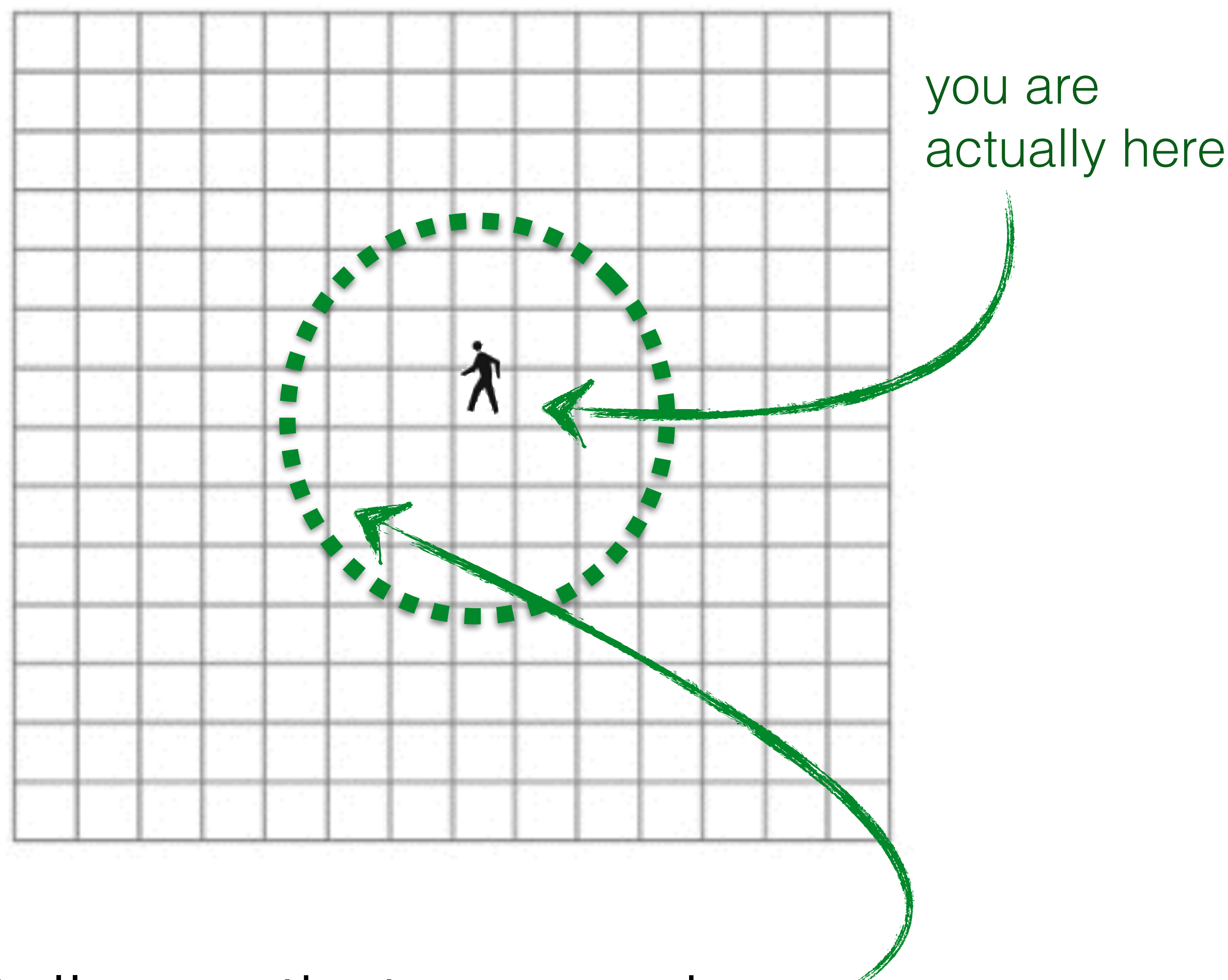
The current observation is usually
most influenced by the current state

$$P(\boldsymbol{E}_t|\boldsymbol{X}_t)$$

(this relationship is called the **observation** model)

Can you think of an observation of a state?

For example, GPS is a noisy observation of location.

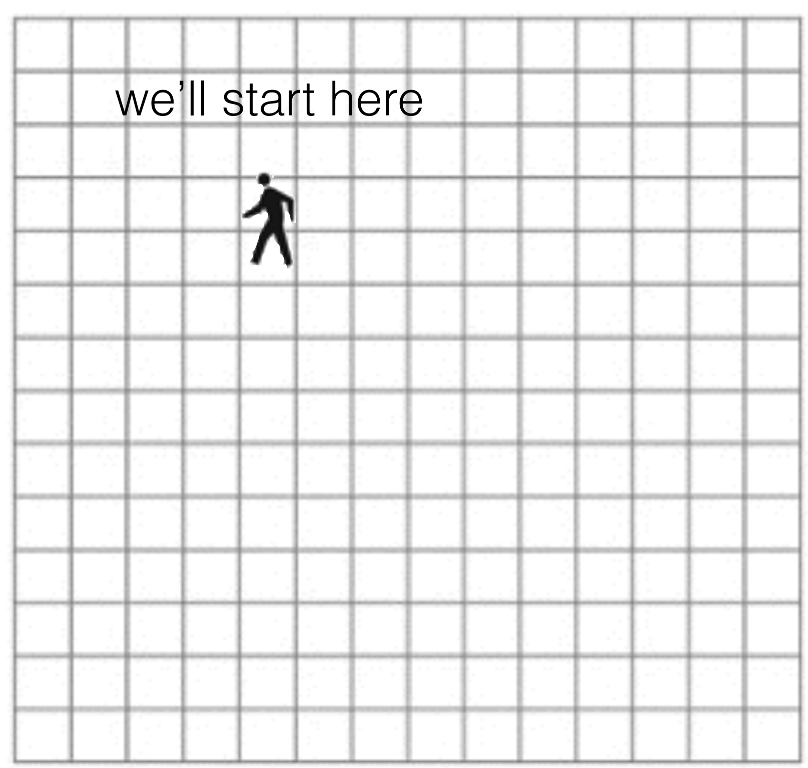


But GPS tells you that you are here
with probability $P(\boldsymbol{E}_t|\boldsymbol{X}_t)$

Reduction 3. Prior State Assumption:

‘we know where the process (probably) starts’

$$X_0$$



Applying these assumptions,
we can decompose the joint probability:

$$P(\boldsymbol{X}_0 \boldsymbol{X}_1, \ldots, \boldsymbol{X}_T, \boldsymbol{E}_1 \boldsymbol{E}_1, \ldots, \boldsymbol{E}_T) = P(\boldsymbol{X}_0) \prod_{t=1}^{T} P(\boldsymbol{X}_t | \boldsymbol{X}_{t-1}) P(\boldsymbol{E}_t | \boldsymbol{X}_t)$$

Stationary process assumption:

only have to store ____ models

(assuming only a single variable for state and observation)

Markov assumption:

This is a model of order ___

We have significantly reduced the number of parameters

Joint Probability of a Temporal Sequence

$$P(\boldsymbol{X}_0) \prod_{t=1}^{T} P(\boldsymbol{X}_t | \boldsymbol{X}_{t-1}) P(\boldsymbol{E}_t | \boldsymbol{X}_t)$$

state prior
prior

motion model
transition model

sensor model
observation model

Joint Probability of a Temporal Sequence

$$P(\boldsymbol{X}_0)\prod_{t=1}^{T} P(\boldsymbol{X}_t|\boldsymbol{X}_{t-1})P(\boldsymbol{E}_t|\boldsymbol{X}_t)$$

state prior
prior

motion model
transition model

sensor model
observation model

Joint Distribution for a Dynamic Bayesian Network

specific instances of a DBN
covered in this class

Hidden Markov Model

(typically taught as discrete but not necessarily)

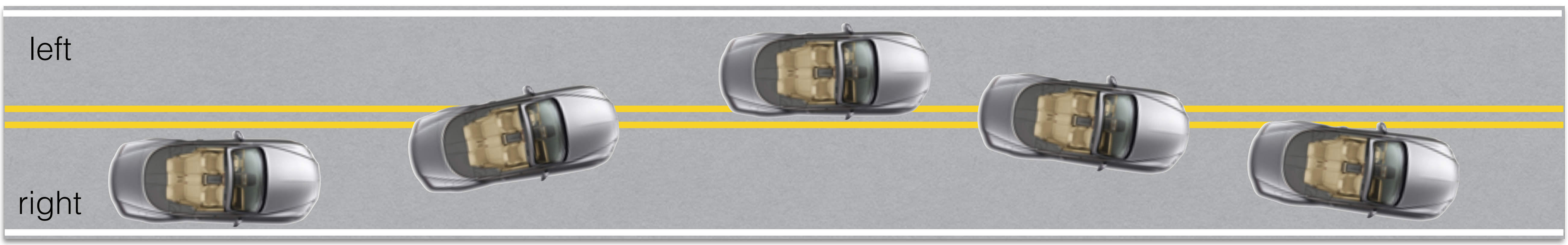
Kalman Filter

(Gaussian motion model, prior and observation model)

Hidden Markov Model

Hidden Markov Model example

'In the trunk of a car of a sleepy driver' model



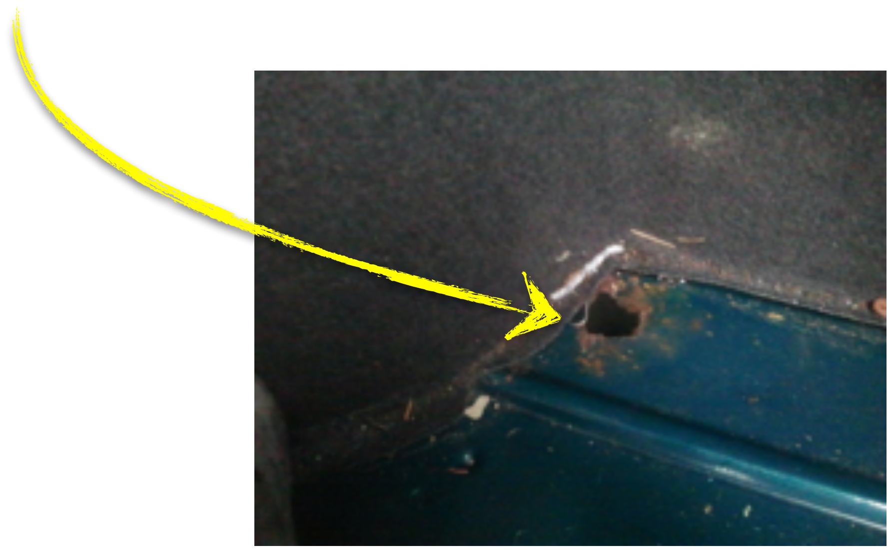
binary random variable (left lane or right lane)

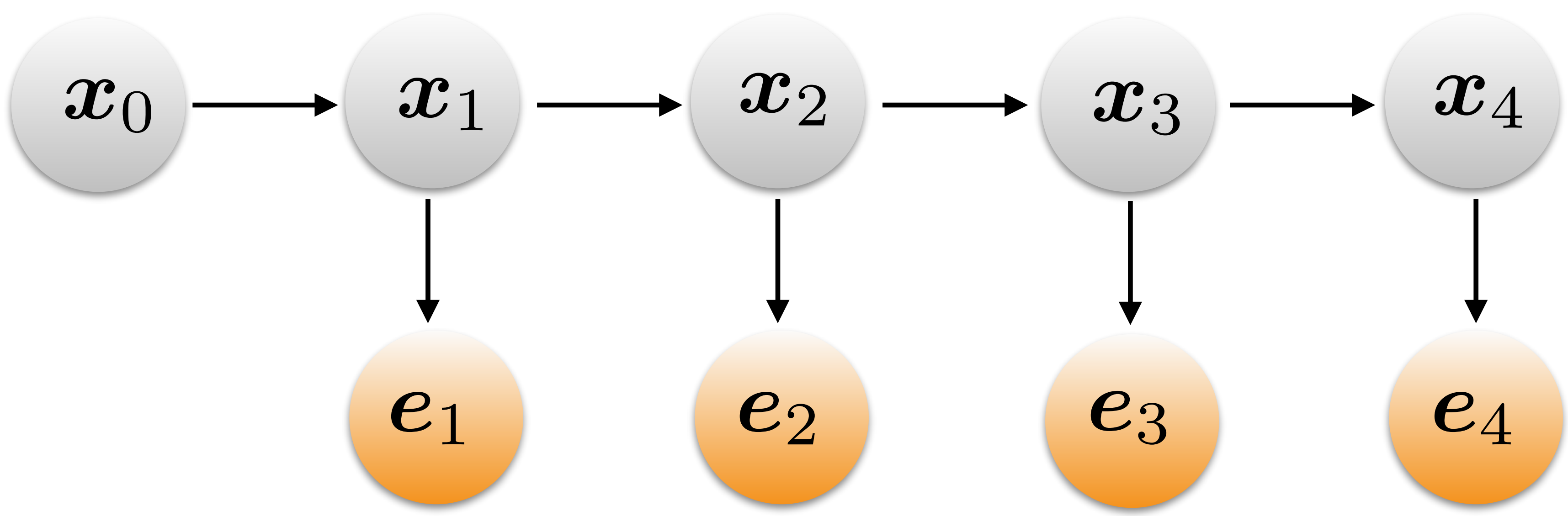
$$\boldsymbol{x}_0 \rightarrow \boldsymbol{x}_1 \rightarrow \boldsymbol{x}_2 \rightarrow \boldsymbol{x}_3 \rightarrow \boldsymbol{x}_4$$

$$\boldsymbol{x} = \{x_{\text{left}}, x_{\text{right}}\}$$

two state world!

From a hole in the car you can see the ground





binary random variable (road is yellow or road is gray)

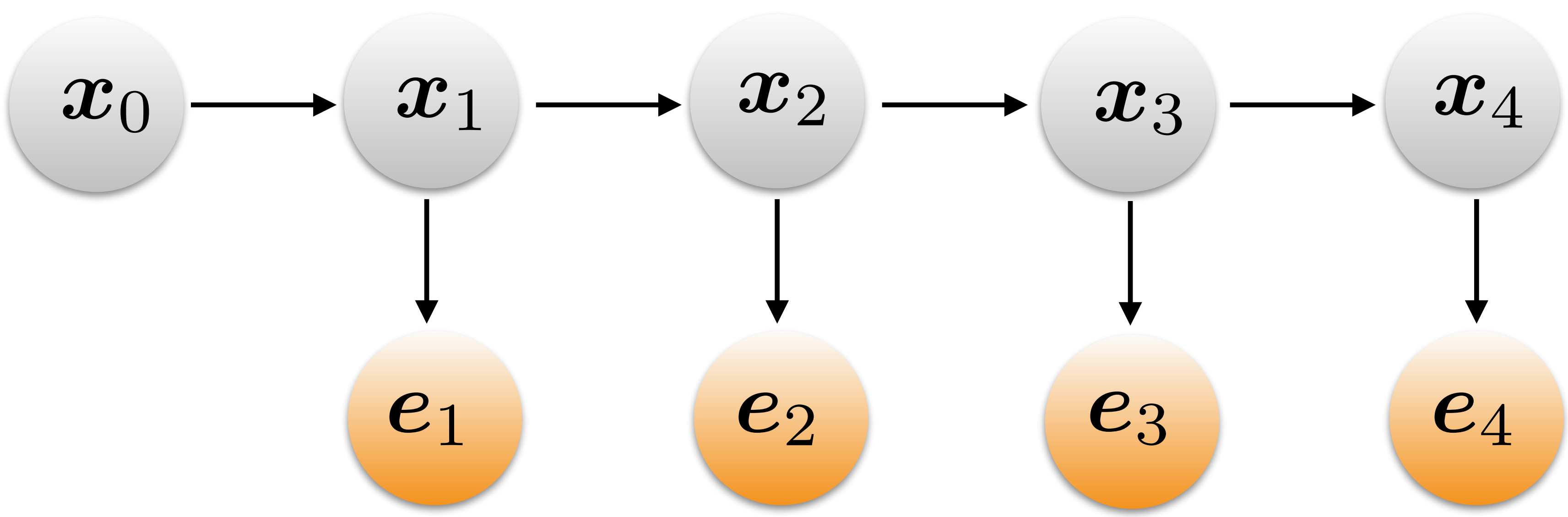
$$\boldsymbol{e} = \{e_{\text{gray}}, e_{\text{yellow}}\}$$

Nice visualization here: http://setosa.io/ev/markov-chains/?utm_content=buffer5e504

	x_{left}	x_{right}
$P(\boldsymbol{x}_0)$	0.5	0.5

$P(\boldsymbol{x}_t \mid \boldsymbol{x}_{t-1})$	x_{left}	x_{right}
x_{left}	0.7	0.3
x_{right}	0.3	0.7

What needs to sum to one?



What's the probability of staying in the left lane if I'm in the left lane?

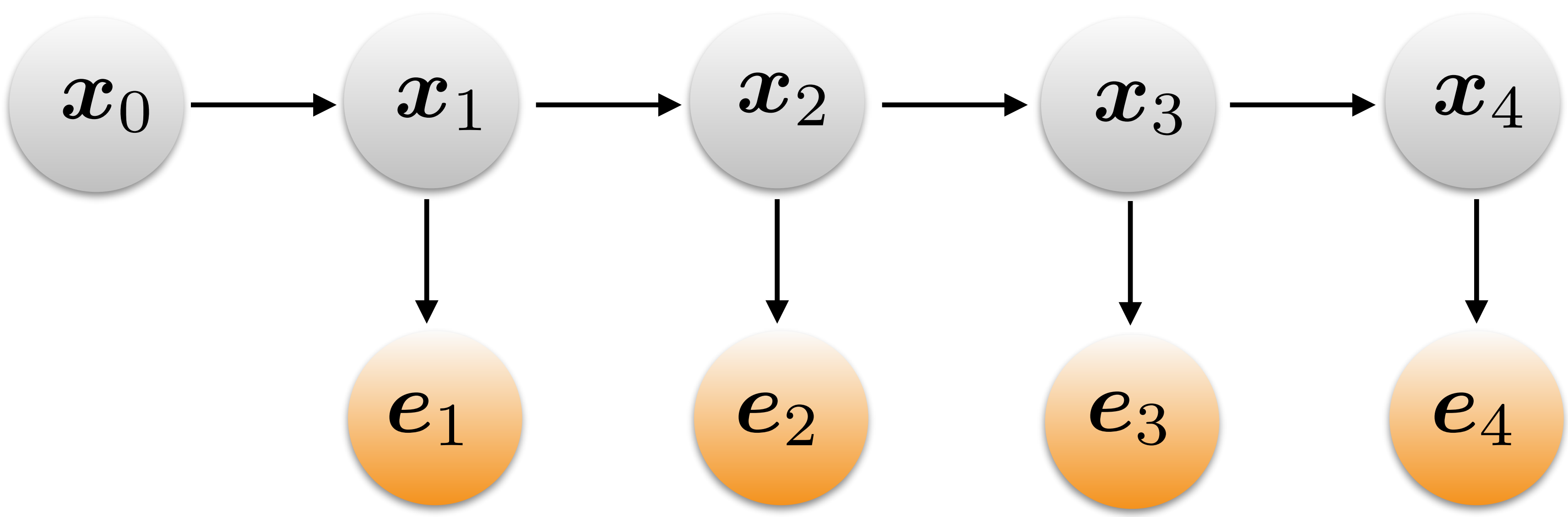
$P(\boldsymbol{e}_t \mid \boldsymbol{x}_t)$	x_{left}	x_{right}
e_{yellow}	0.9	0.2
e_{gray}	0.1	0.8

What lane am I in if I see yellow?

visualization of the motion model

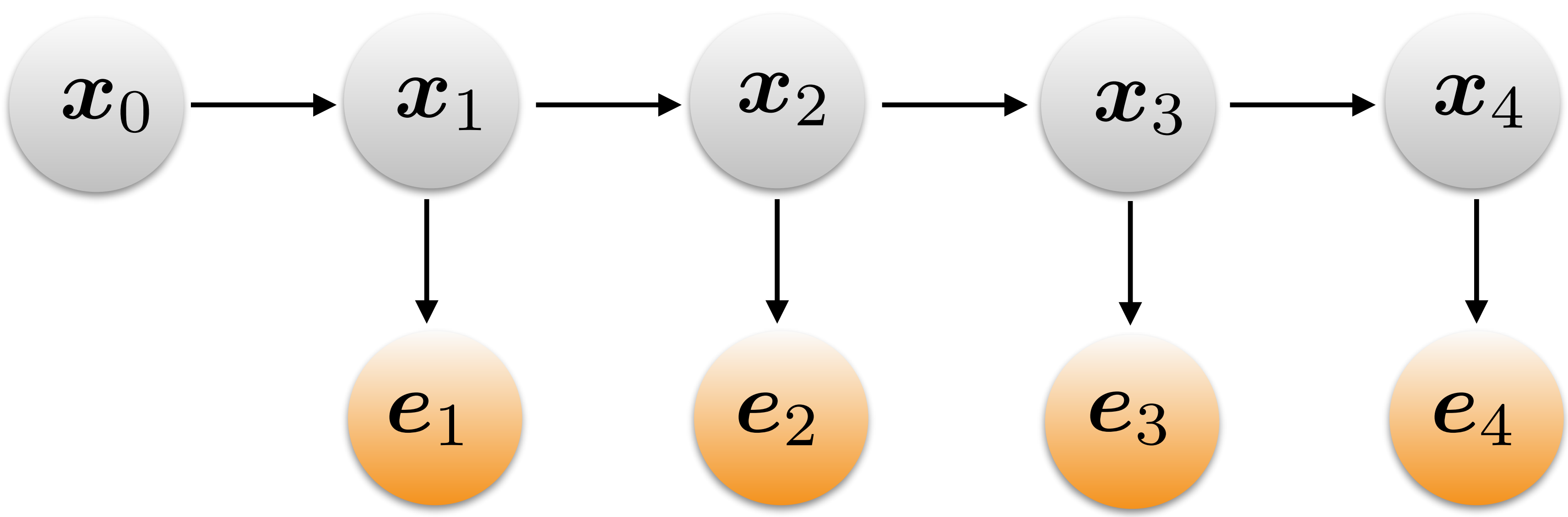
$P(\boldsymbol{x}_t \mid \boldsymbol{x}_{t-1})$	$x_{t-1} = R$	$x_{t-1} = S$
$x_t = R$	0.9	0.1
$x_t = S$	0.1	0.9

Is the stationary assumption true?



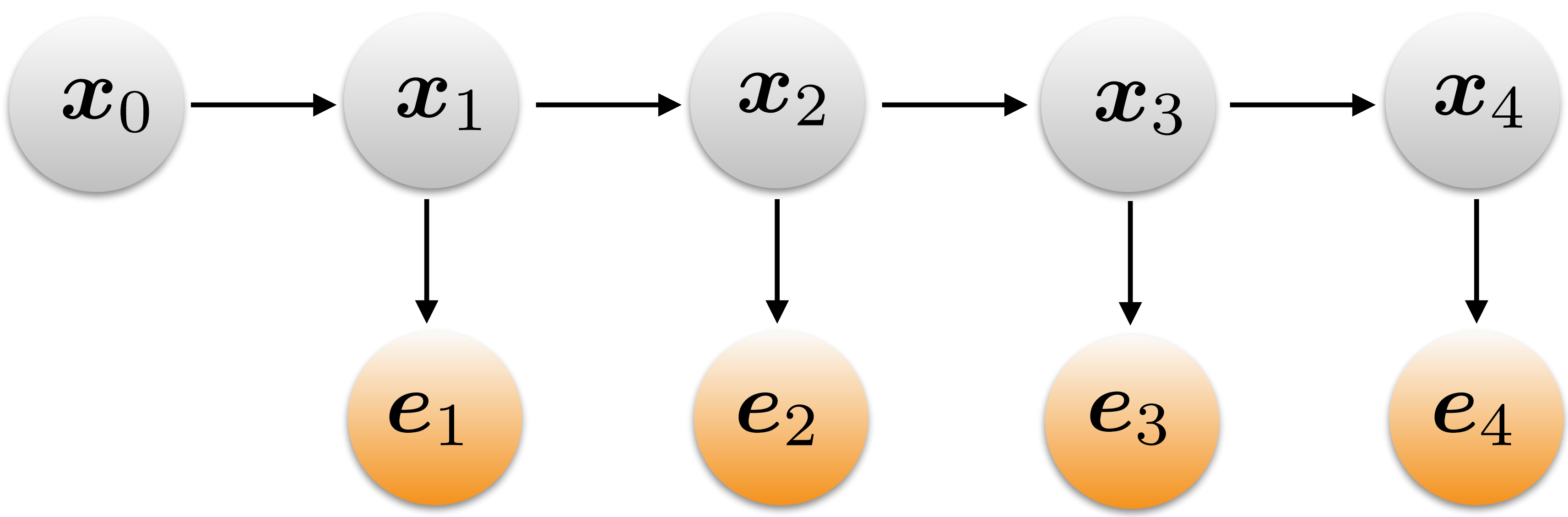
Is the stationary assumption true?

visibility at night?
visibility after a day in the car?
still swerving after one day of driving?



Is the stationary assumption true?

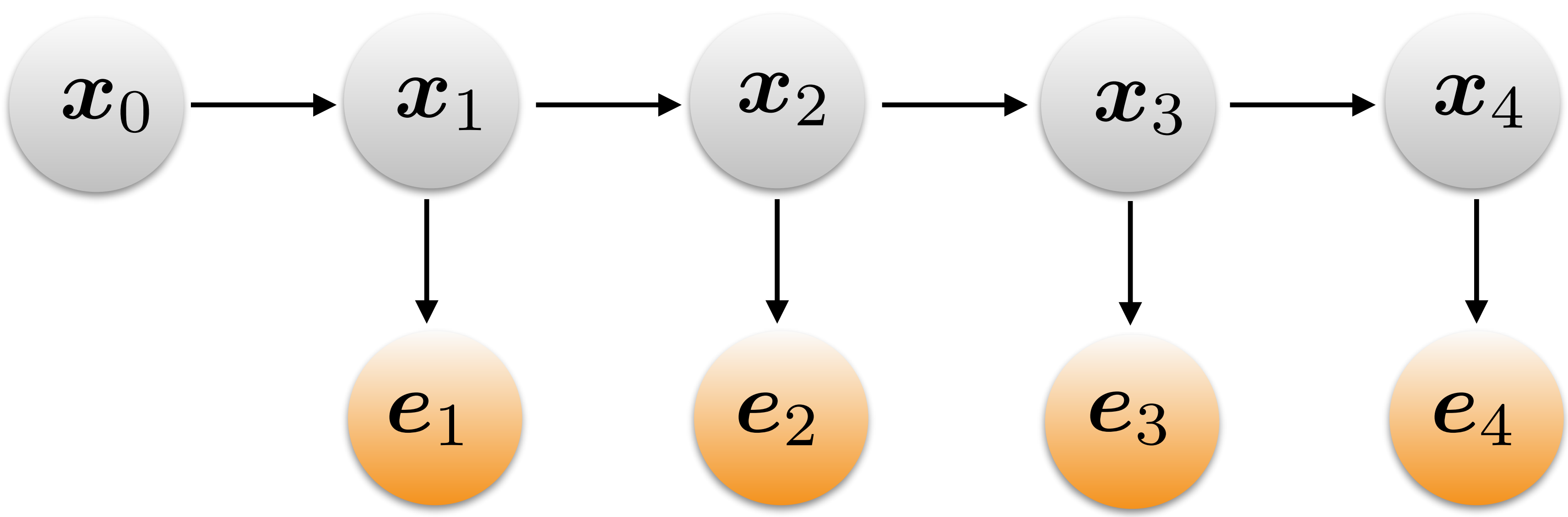
visibility at night?
visibility after a day in the car?
still swerving after one day of driving?



Is the Markov assumption true?

Is the stationary assumption true?

visibility at night?
visibility after a day in the car?
still swerving after one day of driving?



Is the Markov assumption true?

what can you learn with higher order models?
what if you have been in the same lane for the last hour?

In general, assumptions are not correct but they simplify the problem and work most of the time when designed appropriately