Guided Tour of Machine Learning in Finance

Week 4: Reinforcement Learning

4-1-2-Latent Variable models for sequences

Igor Halperin

NYU Tandon School of Engineering, 2017

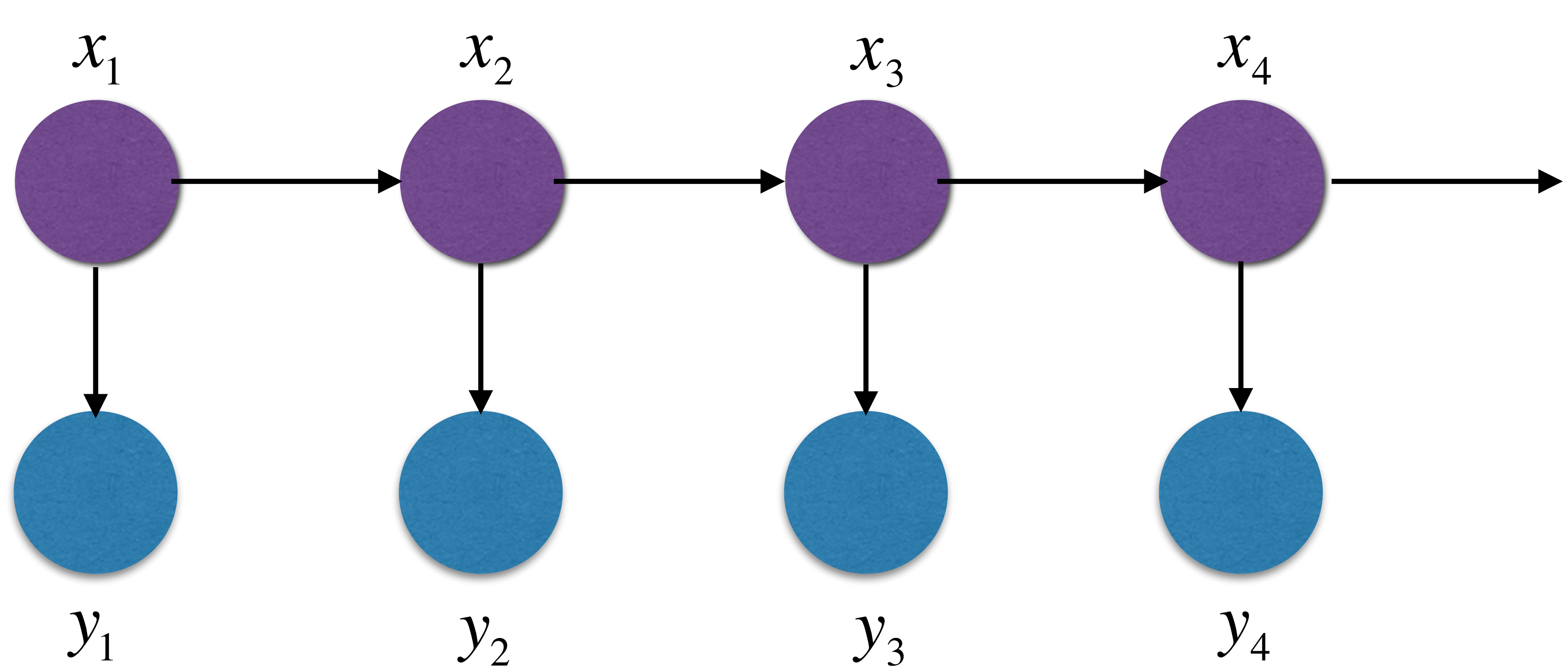
How to model sequential data

The last: build a probabilistic model for the observable signal $y = \left\{ y^{(t)} \right\}_{t=1}^{T}$

$$p\left(y^{(1)}, y^{(2)}, y^{(3)}, \ldots, y^{(T)}\right) = \prod_{n=1}^{T} p\left(y^{(n)} \mid y^{(1)}, y^{(2)}, \ldots, y^{(n-1)}\right)$$

Possible modeling approaches for $p\left(y^{(t)} \mid y^{(t-1)}, y^{(t-2)}, \ldots\right)$:

1. **Markov models:**
2. **Dynamic hidden (latent) variables models**, with an unobservable state



How to model sequential data

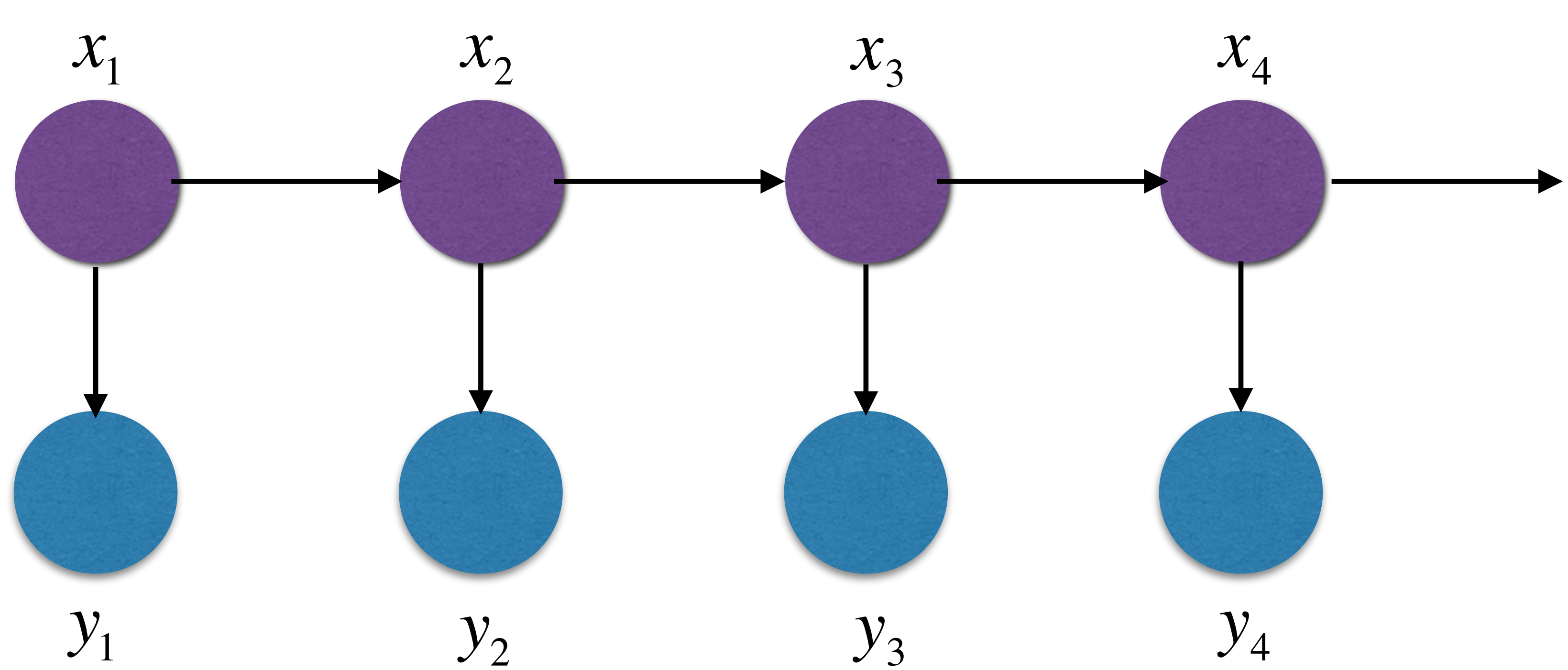
The last: build a probabilistic model for the observable signal $y = \left\{ y^{(t)} \right\}_{t=1}^{T}$

$$p\left(y^{(1)}, y^{(2)}, y^{(3)}, \ldots, y^{(T)}\right) = \prod_{n=1}^{T} p\left(y^{(n)} \mid y^{(1)}, y^{(2)}, \ldots, y^{(n-1)}\right)$$

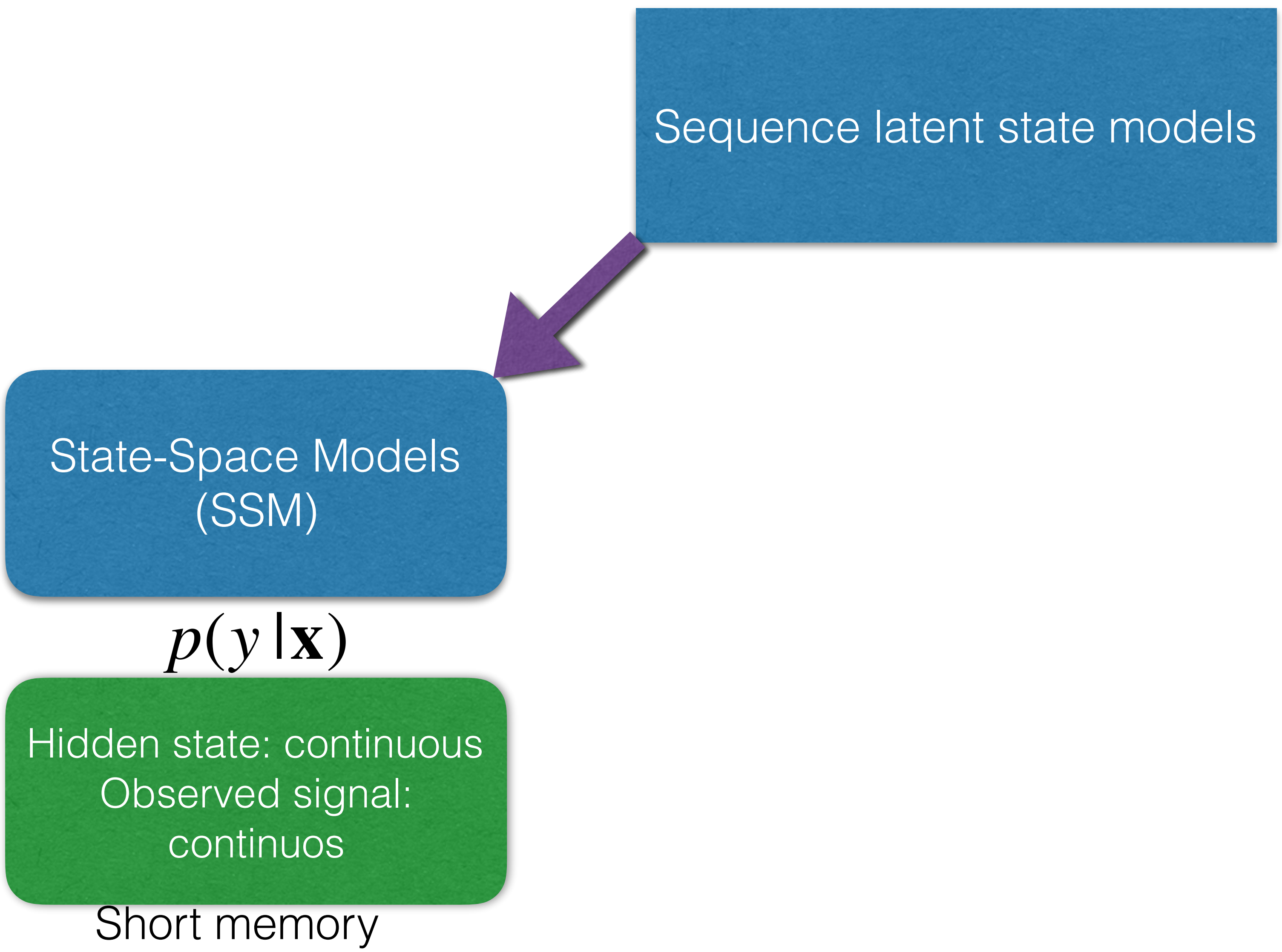
Possible modeling approaches for $p\left(y^{(t)} \mid y^{(t-1)}, y^{(t-2)}, \ldots\right)$:

1. **Markov models:**
2. **Dynamic hidden (latent) variables models**, with an unobservable state $x^{(t)}$

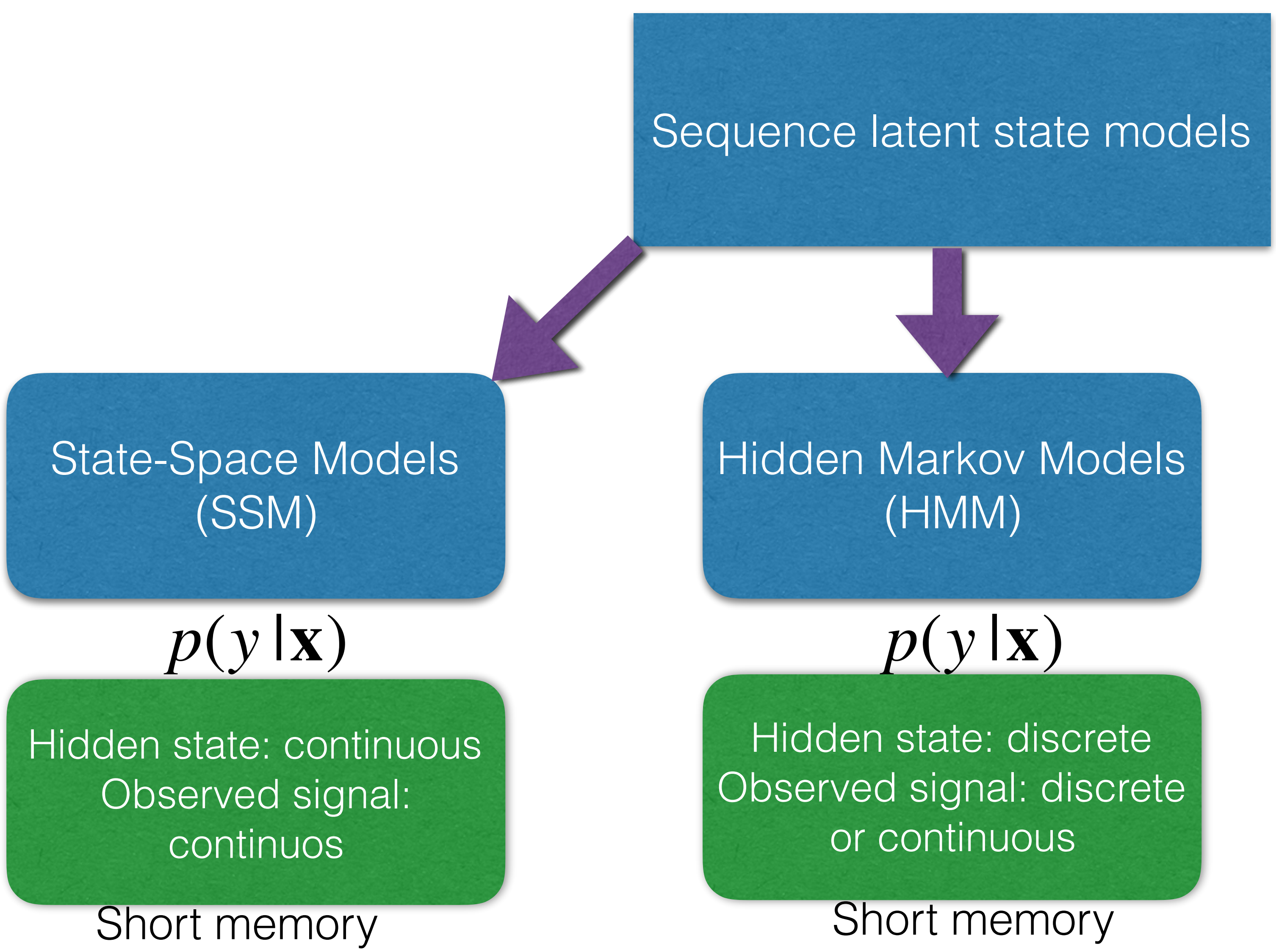
- The hidden state $x^{(t)}$ captures the dynamics of the system, filters noise out
- Used as a conditioning variable for predictions $p\left(y^{(t)} \mid y^{(t-1)}\right) \rightarrow p\left(y^{(t)} \mid x^{(t)}\right)$.



Sequence latent state models

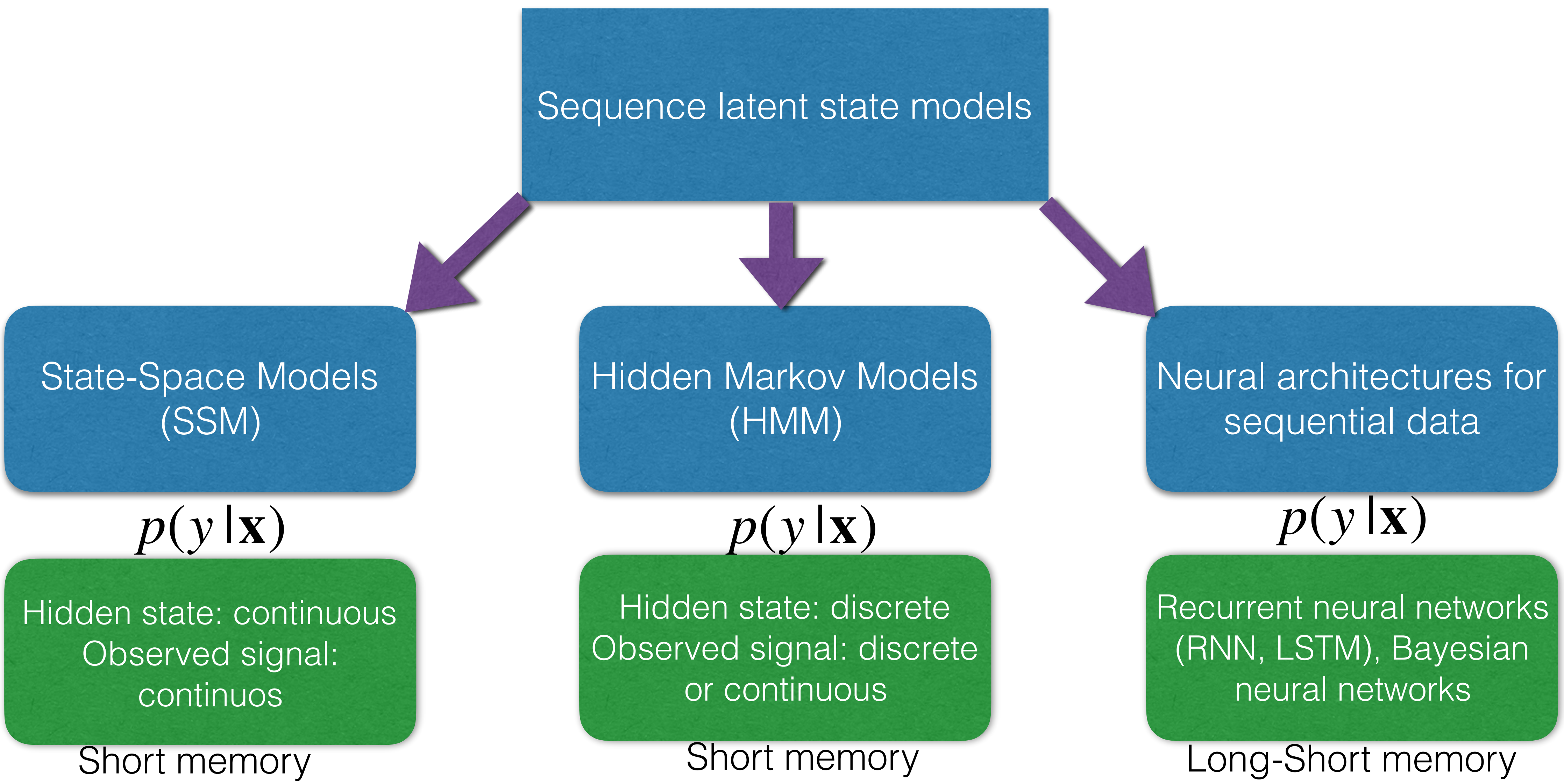


Sequence latent state models



Sequence latent state models

Parametric (SSM, HMM) vs non-parametric (neural) of sequence modeling with a hidden state $p(y\,|\,\mathbf{x})$



Control question

Select all correct answers

1. State-Space models measure the fraction of total space occupied a given state of a system, then assign probabilities to all states proportional to these fractions.
2. The Hidden Markov Models (HMM) are named so because in these models the Markov dynamics is hidden (masked by observational noise), therefore they model the dynamics as non-Markov.
3. For State-Space Models (SSM), both the hidden and observed states are continuous.
4. For Hidden Markov Models (HMM), the hidden state is discrete, while the observed state can be either discrete or continuous
5. For Neural models of sequential data, both the hidden and observed states can be either discrete or continuous, as long as they are defined as TensorFlow constants on the graph.
6. The abbreviation LSTM means Long Term Capital Management.

Correct answers: 3, 4

Estimation of Dynamic Hidden Variable Models

Observable N-dimensional data: $y_{1:T} = y^{(1)}, y^{(2)}, \ldots, y^{(T)}$
Hidden state sequence: $x_{1:T} = x^{(1)}, x^{(2)}, \ldots, x^{(T)}$
The dynamics is first-order Markov in the hidden state (either for a SSM or HMM):

$$p(x_{1:T}, y_{1:T} \mid \theta) = \prod_{t=1}^{T} p(x_t \mid x_{t-1}, \theta)\, p(y_t \mid x_t, \theta)$$

- Hidden states x have first-order Markov dynamics encoded in $p(x_t \mid x_{t-1}, \theta)$
- Observations are generated from hidden states according to $p(y_t \mid x_t, \theta)$

Log-Likelihood of data**:**

$$\log LL = \log \int \prod_{t} p(y_{t+1} \mid x_{1:t}, y_t, \theta)\, dx_{1:t}$$

This can be estimated using the **EM algorithm**.