Methods For Working With Time Series: Hidden Markov Models & More

Hunter Glanz

California Polytechnic State University
San Luis Obispo

February 8, 2019

CAL POLY
SAN LUIS OBISPO

Outline

1. Introduction to Time Series
2. Traditional Time Series Analysis
3. Introduction to Hidden Markov Models

CAL POLY
SAN LUIS OBISPO

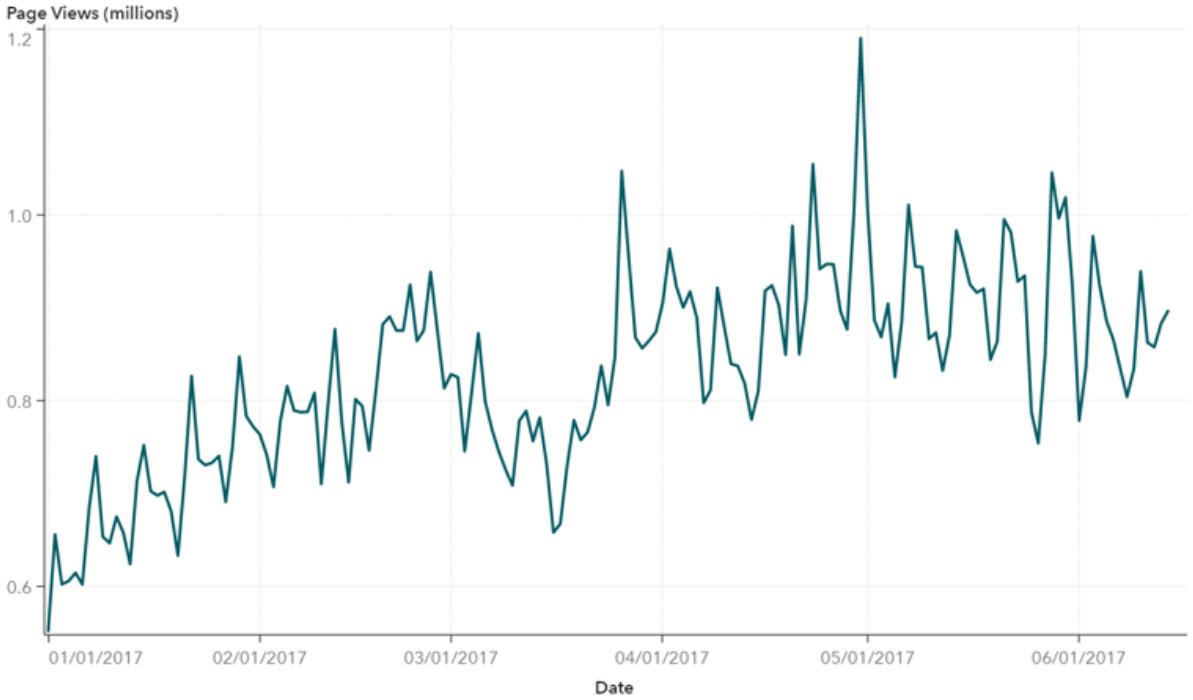
Outline

1. Introduction to Time Series
2. Traditional Time Series Analysis
3. Introduction to Hidden Markov Models

CAL POLY
SAN LUIS OBISPO

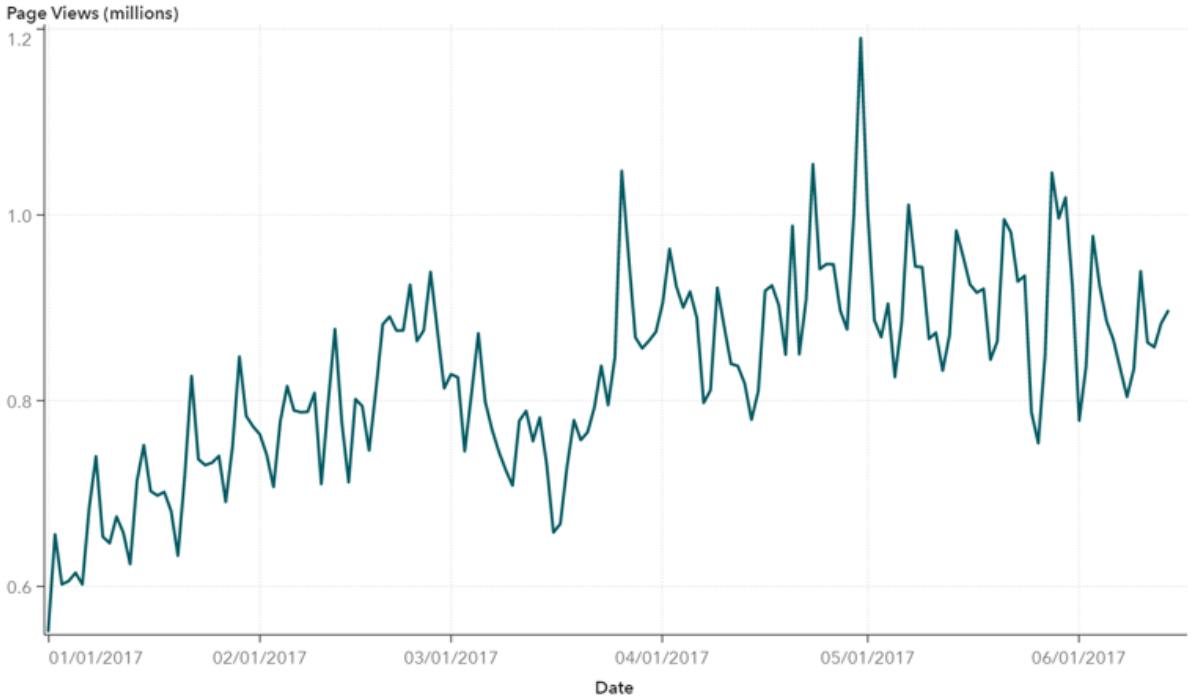
What is a *time series*?

What is a *time series*?



CAL POLY
SAN LUIS OBISPO

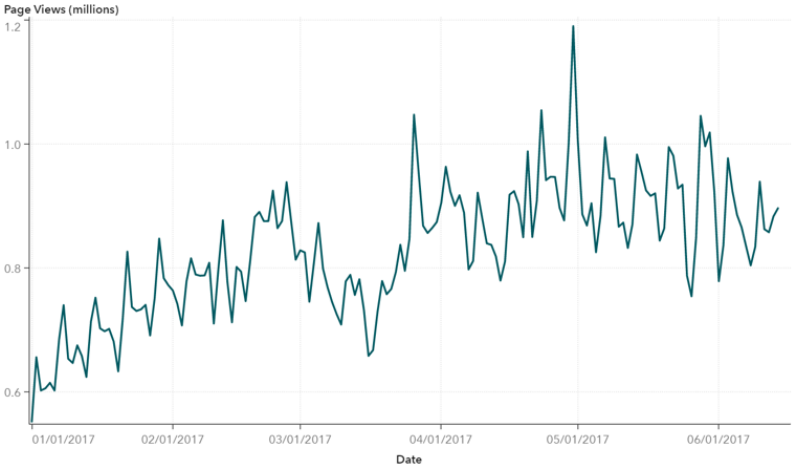
What is a *time series*?



- **time series**: a series of data points indexed in time order; most commonly taken at successive equally spaced points in time.

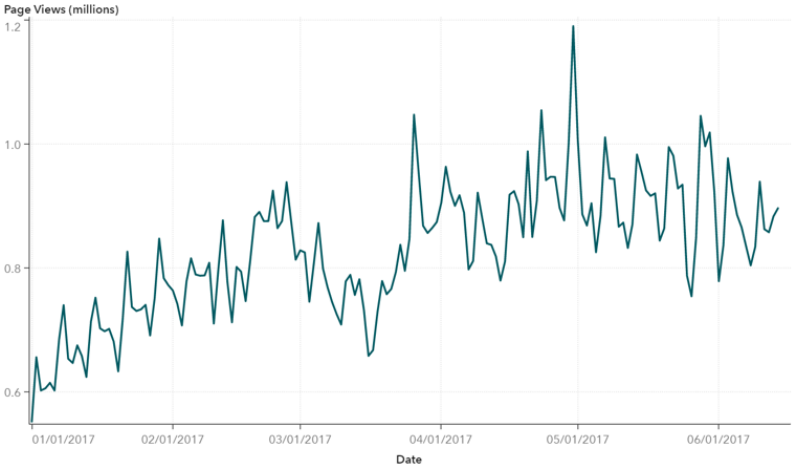
CAL POLY
SAN LUIS OBISPO

What Are We Interested In?



CAL POLY
SAN LUIS OBISPO

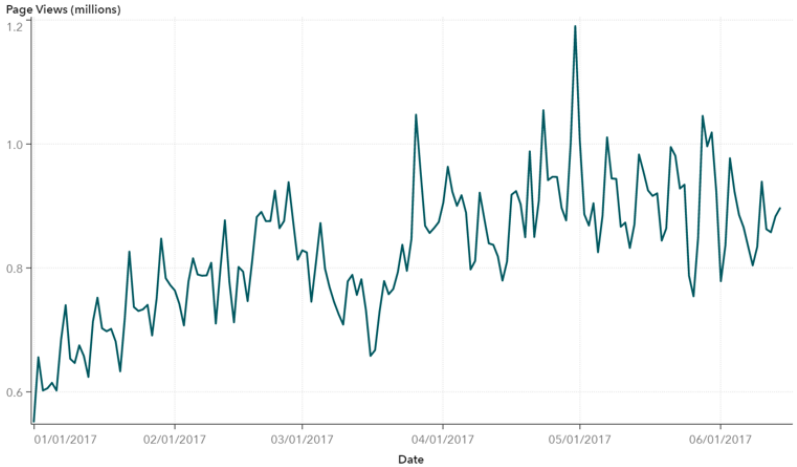
What Are We Interested In?



- Descriptive/Exploratory

CAL POLY
SAN LUIS OBISPO

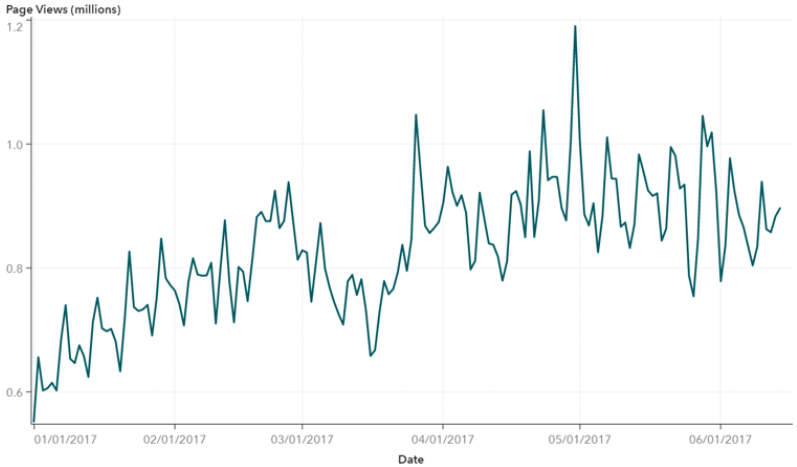
What Are We Interested In?



- Descriptive/Exploratory
- Curve Fitting/Function Approximation

CAL POLY
SAN LUIS OBISPO

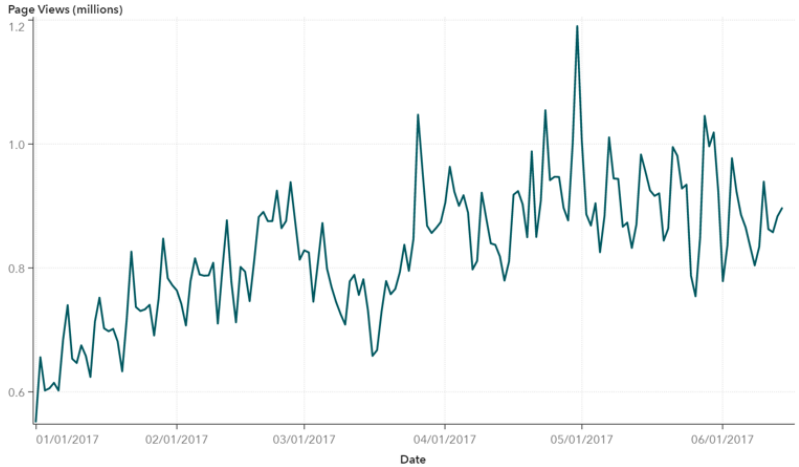
What Are We Interested In?

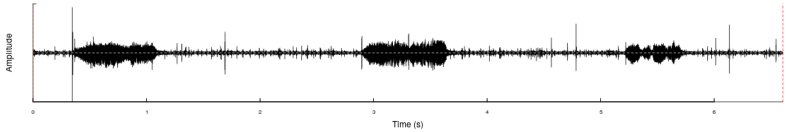


- Descriptive/Exploratory
- Curve Fitting/Function Approximation
- **Prediction/Forecasting**

CAL POLY
SAN LUIS OBISPO

What Are We Interested In?





- Descriptive/Exploratory
- Curve Fitting/Function Approximation
- **Prediction/Forecasting**
- **Segmentation/Classification**

CAL POLY
SAN LUIS OBISPO

Outline

1. Introduction to Time Series
2. Traditional Time Series Analysis
3. Introduction to Hidden Markov Models

CAL POLY
SAN LUIS OBISPO

STAT 416: Statistical Analysis of Time Series

Analysis and forecasting of a single quantitative variable (time series)

CAL POLY
SAN LUIS OBISPO

STAT 416: Statistical Analysis of Time Series

Analysis and forecasting of a single quantitative variable (time series)

- Autocorrelation
- Autoregressive (AR) models
- Moving Average (MA) models
- ARMA & ARIMA models

CAL POLY
SAN LUIS OBISPO

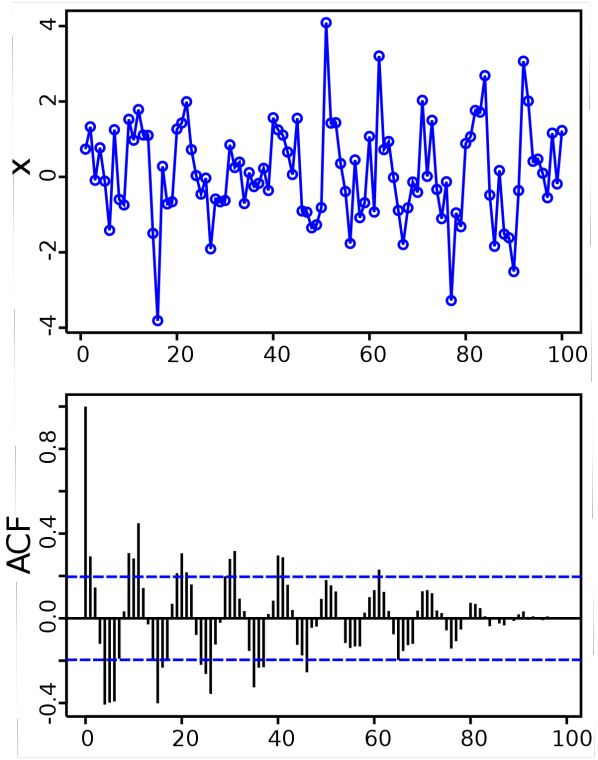
Autocorrelation

Autocorrelation

Correlation of a signal with a delayed copy of itself as a function of the delay

Autocorrelation

Correlation of a signal with a delayed copy of itself as a function of the delay

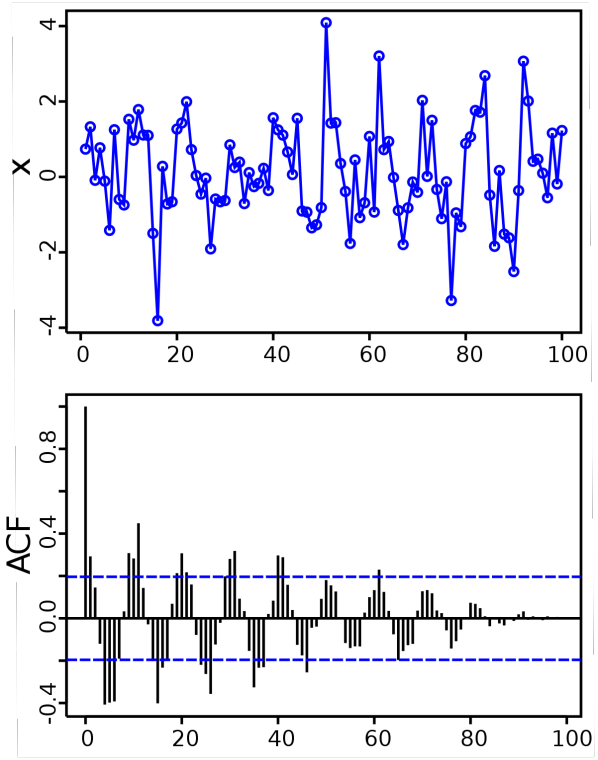


- **autocorrelation**: similarity between observations as a function of the time lag between them

CAL POLY
SAN LUIS OBISPO

Autocorrelation

Correlation of a signal with a delayed copy of itself as a function of the delay



- **autocorrelation**: similarity between observations as a function of the time lag between them
- What could you conclude from the graph of the ACF?

Autoregressive Models

Autoregressive Models

- Autoregressive model of order p; $AR(p)$.

$$X_t = c + \sum_{i=1}^{p} \varphi_i X_{t-i} + \varepsilon_t$$

CAL POLY
SAN LUIS OBISPO

Autoregressive Models

- Autoregressive model of order p; $AR(p)$.

$$X_t = c + \sum_{i=1}^{p} \varphi_i X_{t-i} + \varepsilon_t$$

- How do we choose p?
- How do we estimate the φ coefficients?

CAL POLY
SAN LUIS OBISPO

Moving Average Models

Moving Average Models

- Moving Average model of order q; $MA(q)$.

$$X_t = \mu + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \cdots + \theta \varepsilon_{t-q}$$

Moving Average Models

- Moving Average model of order q; $MA(q)$.

$$X_t = \mu + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \cdots + \theta \varepsilon_{t-q}$$

- How do we choose q?
- How do we estimate the θ coefficients?

CAL POLY
SAN LUIS OBISPO

ARMA and ARIMA Models

- $ARMA(p, q)$:

$$X_t = c + \varepsilon_t + \sum_{i=1}^{p} \varphi X_{t-i} + \sum_{j=1}^{q} \theta_j \varepsilon_{t-j}$$

CAL POLY
SAN LUIS OBISPO

ARMA and ARIMA Models

- $ARMA(p, q)$:

$$X_t = c + \varepsilon_t + \sum_{i=1}^{p} \varphi X_{t-i} + \sum_{j=1}^{q} \theta_j \varepsilon_{t-j}$$

- $ARIMA(p, d, q)$:
  - p is the order of the AR part of the model
  - q is the order of the MA part of the model
  - d is the degree of differencing of the data values

CAL POLY
SAN LUIS OBISPO

And Beyond!

- Other methods:
  - spectral analysis
  - wavelet analysis
  - signal processing
  - statistical and machine learning methods

CAL POLY
SAN LUIS OBISPO

And Beyond!

- Other methods:
  - spectral analysis
  - wavelet analysis
  - signal processing
  - statistical and machine learning methods
- Python Implementations
  - **Statsmodels**
  - PyFlux
  - PyMC3

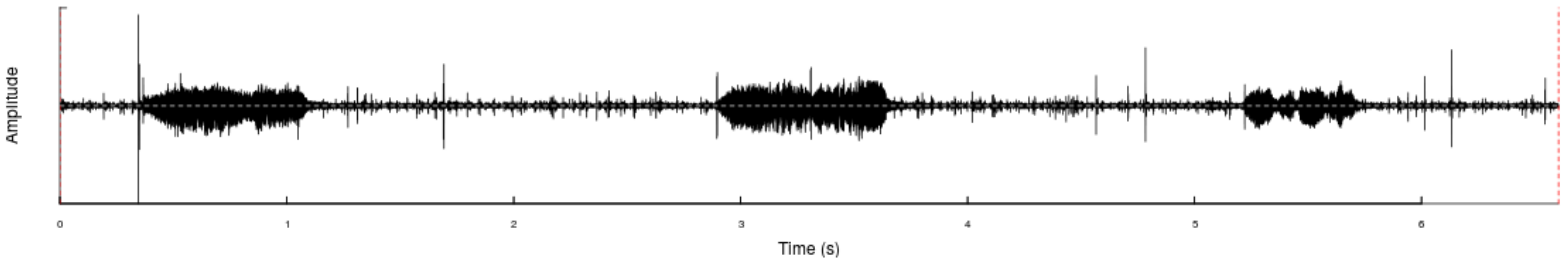
CAL POLY
SAN LUIS OBISPO

Outline

1. Introduction to Time Series
2. Traditional Time Series Analysis
3. Introduction to Hidden Markov Models

CAL POLY
SAN LUIS OBISPO

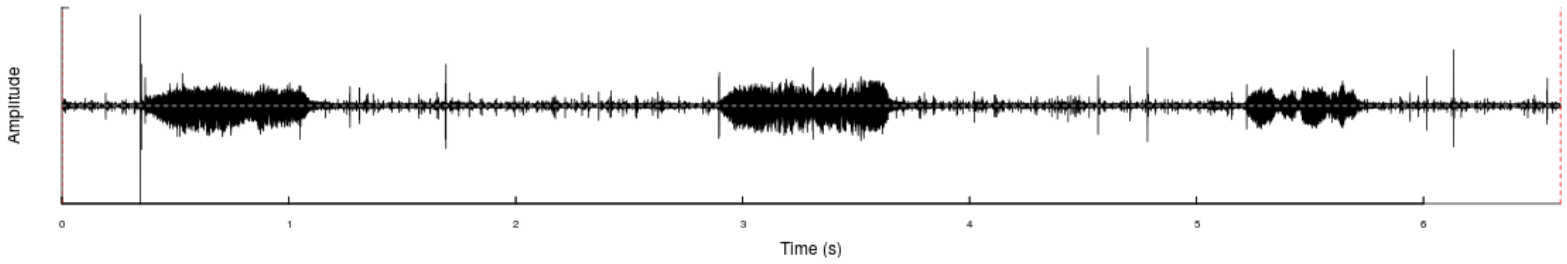
Segmentation and Classification



- What if we're not interested in forecasting a quantitative value?

CAL POLY
SAN LUIS OBISPO

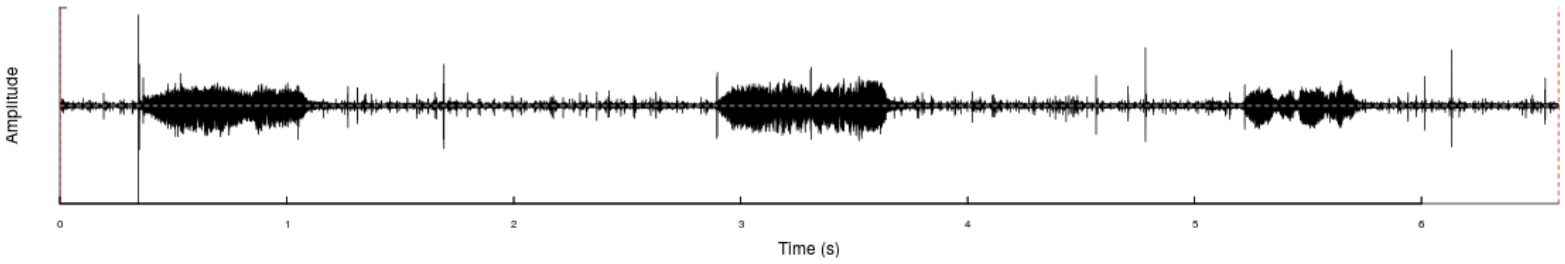
Segmentation and Classification



- What if we're not interested in forecasting a quantitative value?
- Segmentation/Change-point detection

CAL POLY
SAN LUIS OBISPO

Segmentation and Classification



- What if we're not interested in forecasting a quantitative value?
- Segmentation/Change-point detection
- Segmentation $\Longrightarrow$ Classification

CAL POLY
SAN LUIS OBISPO

Discrete Inference

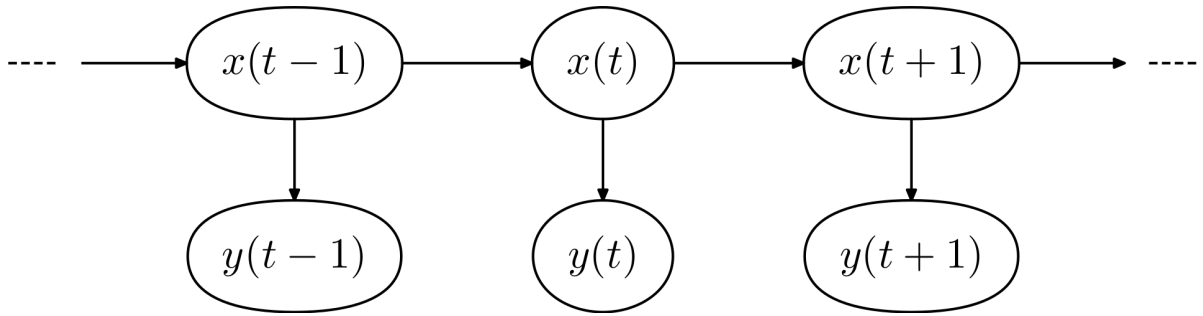
Discrete Inference

- **Hidden Markov Models (HMMs)**:
  - Fall under the umbrella of many different types of models
  - Well summarized by the following image:

CAL POLY
SAN LUIS OBISPO

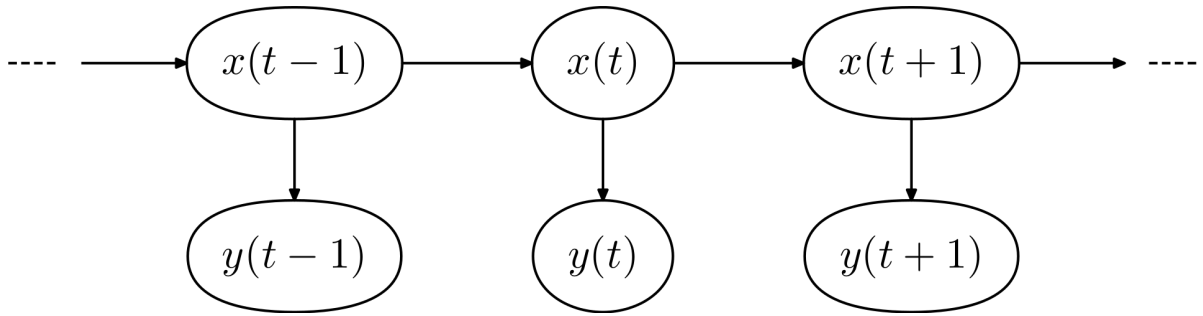
Discrete Inference

- **Hidden Markov Models (HMMs)**:
  - Fall under the umbrella of many different types of models
  - Well summarized by the following image:



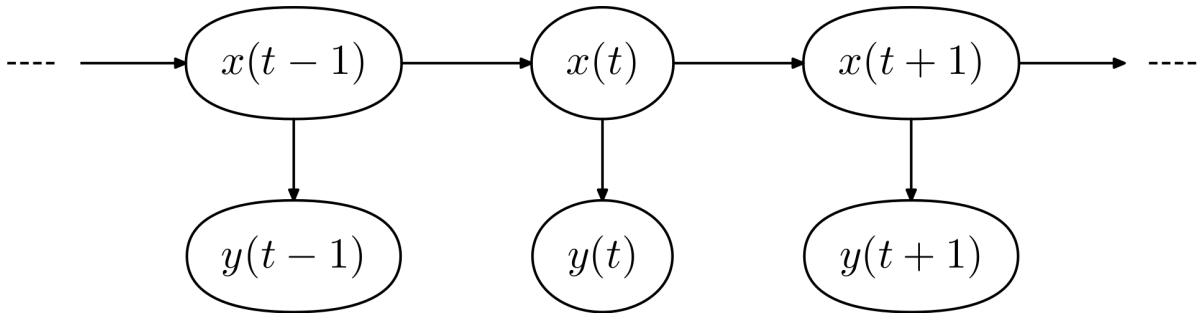
Discrete Inference

- **Hidden Markov Models (HMMs)**:
  - Fall under the umbrella of many different types of models
  - Well summarized by the following image:

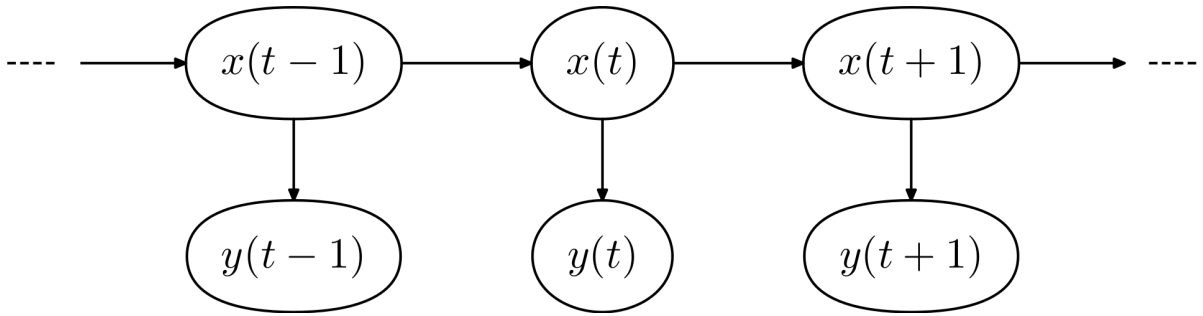


  - Dolphin/Whale calls; Keadle project

Properties of HMMs



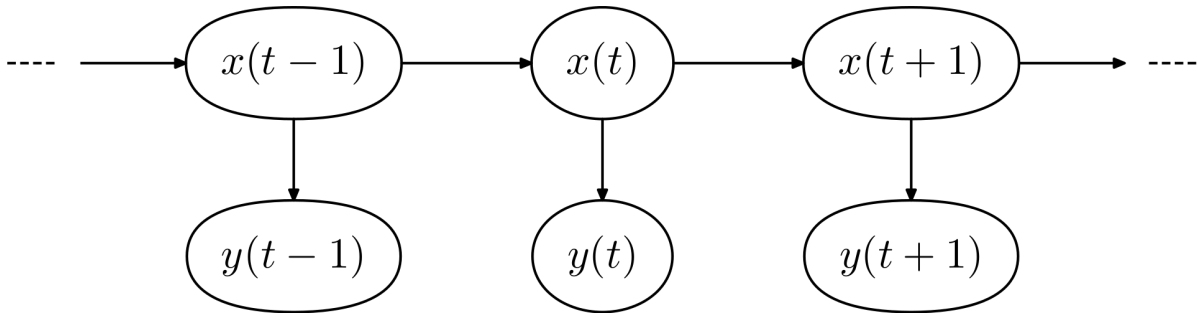
Properties of HMMs



- **Markov property**: Conditional probability distribution of hidden variable, $x(t)$ at time t, depends <u>*only*</u> on the value of the hidden variable $x(t-1)$

CAL POLY
SAN LUIS OBISPO

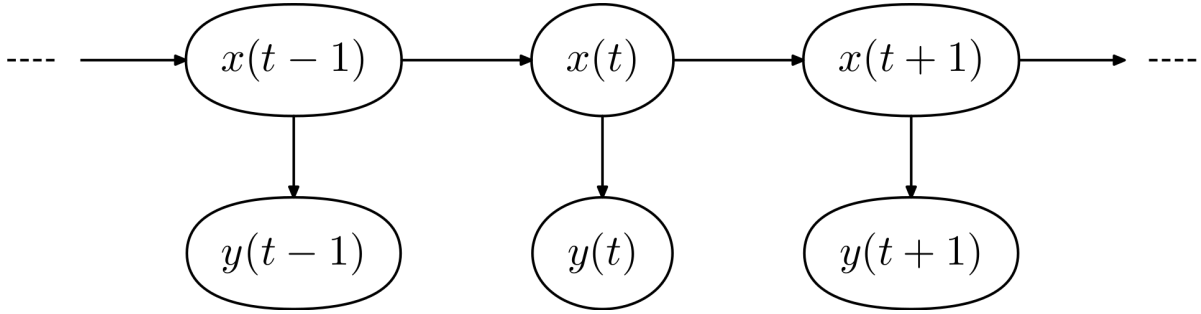
Properties of HMMs



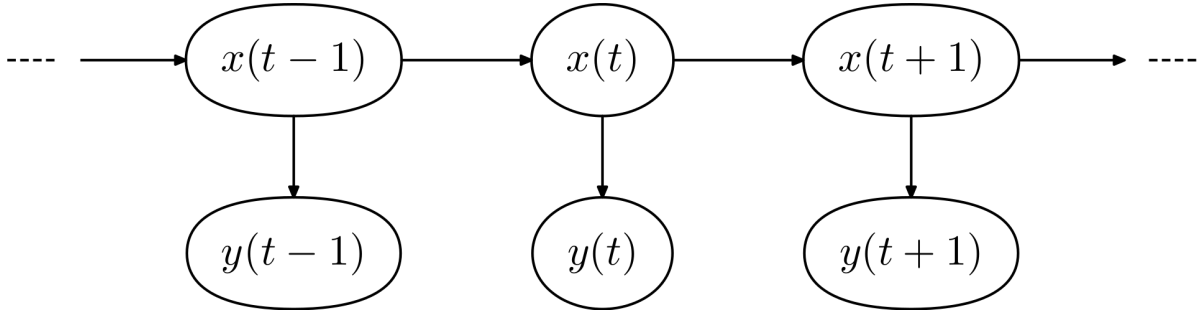
- **Markov property**: Conditional probability distribution of hidden variable, $x(t)$ at time t, depends <u>*only*</u> on the value of the hidden variable $x(t-1)$
- Value of the observed variable $y(t)$ depends <u>*only*</u> on the value of the hidden variable $x(t)$

CAL POLY
SAN LUIS OBISPO

Characterizing HMMs

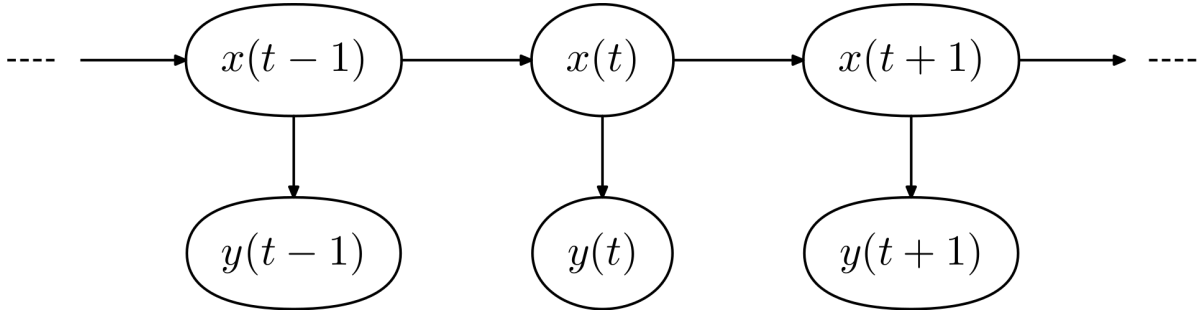


Characterizing HMMs



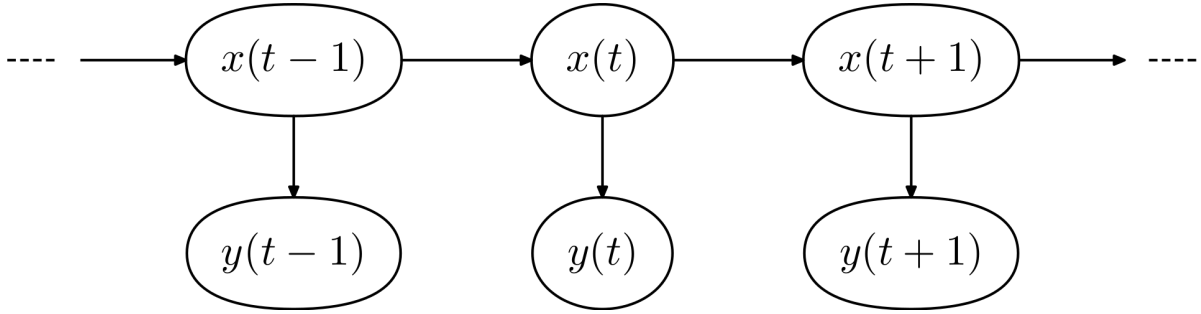
- Number of states

Characterizing HMMs



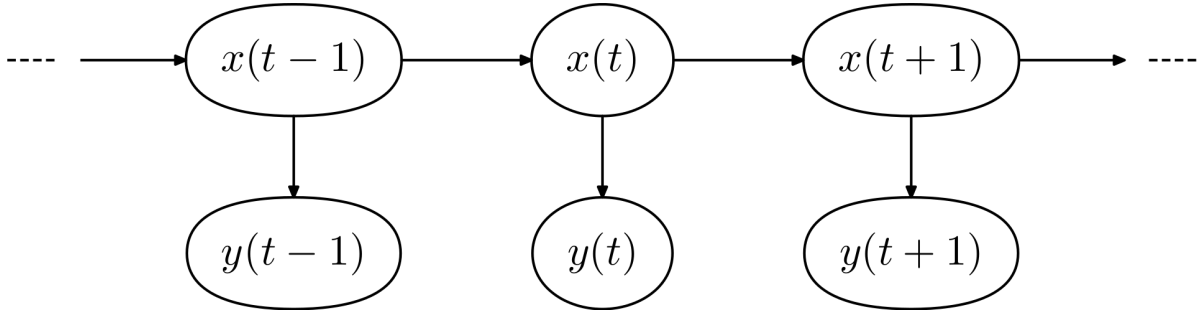
- Number of states
- Number of observations

Characterizing HMMs



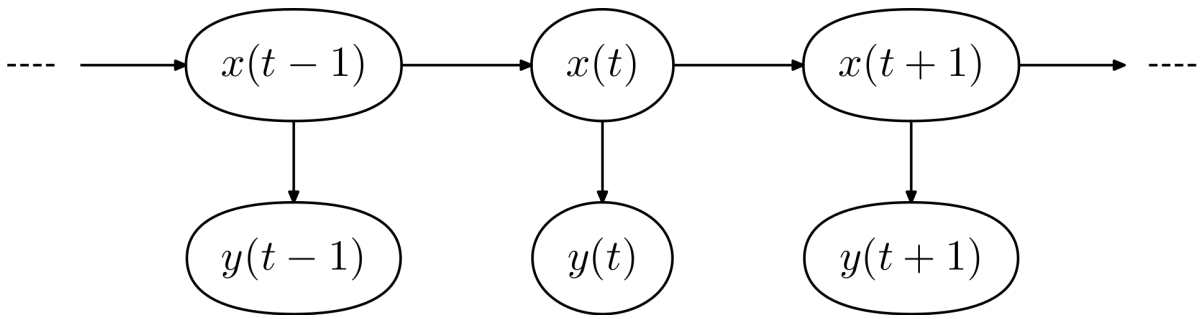
- Number of states
- Number of observations
- Emission probabilities

Characterizing HMMs



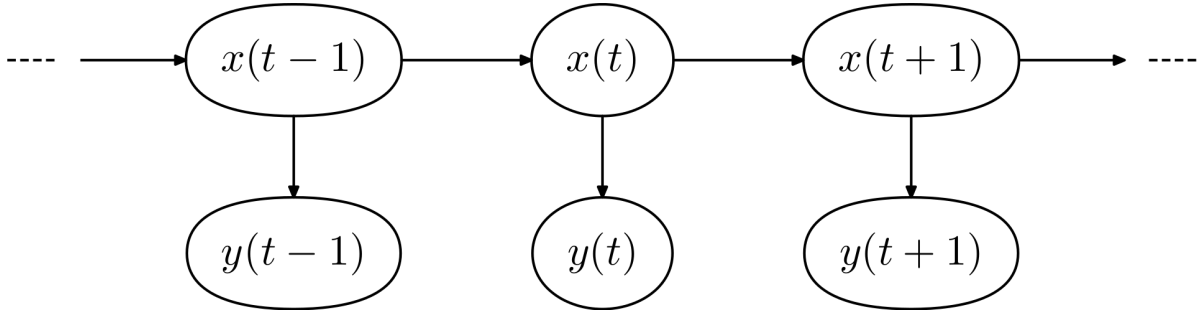
- Number of states
- Number of observations
- Emission probabilities
  - $P(Y_i|X_i)$

Characterizing HMMs



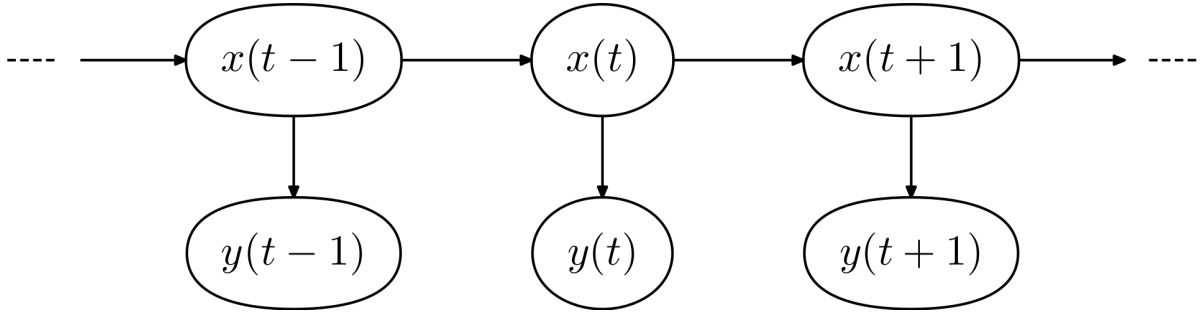
- Number of states
- Number of observations
- Emission probabilities
  - $P(Y_i|X_i)$
- Transition probabilities

Characterizing HMMs



- Number of states
- Number of observations
- Emission probabilities
  - $P(Y_i|X_i)$
- Transition probabilities
  - $P(X_i|X_{i-1})$

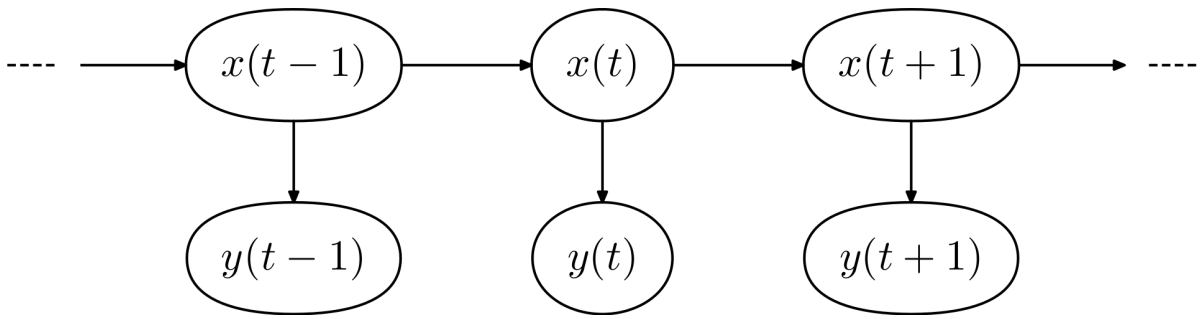
Characterizing HMMs



- Number of states
- Number of observations
- Emission probabilities
  - $P(Y_i|X_i)$
- Transition probabilities
  - $P(X_i|X_{i-1})$
- (Hidden) States

CAL POLY
SAN LUIS OBISPO

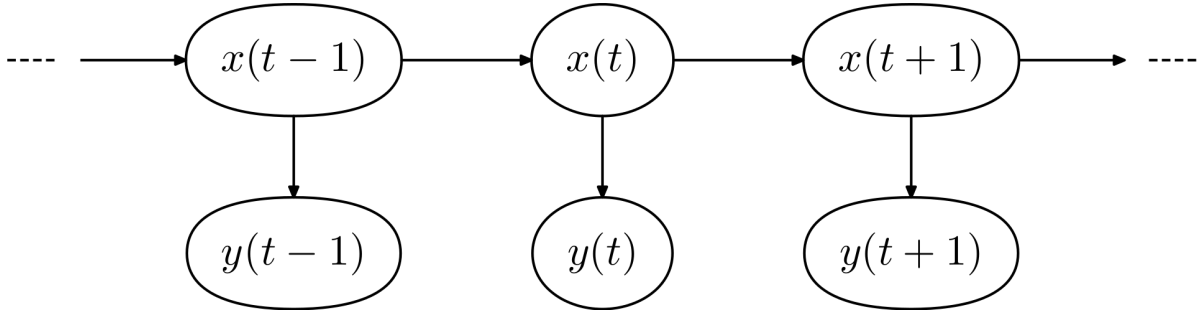
Characterizing HMMs



- Number of states
- Number of observations
- Emission probabilities
  - $P(Y_i|X_i)$
- Transition probabilities
  - $P(X_i|X_{i-1})$
- (Hidden) States
- Observations (data)

CAL POLY
SAN LUIS OBISPO

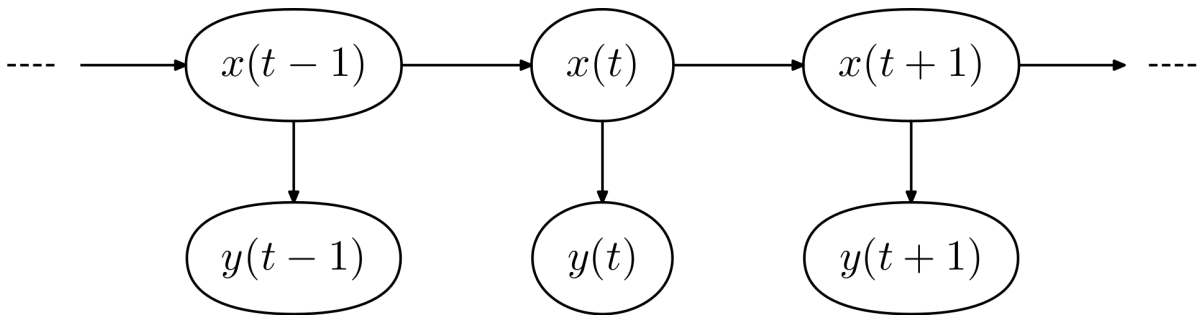
Characterizing HMMs



- Number of states
- Number of observations
- Emission probabilities
  - $P(Y_i|X_i)$
- Transition probabilities
  - $P(X_i|X_{i-1})$
- (Hidden) States
- Observations (data)
- Probability distribution(s)

CAL POLY
SAN LUIS OBISPO

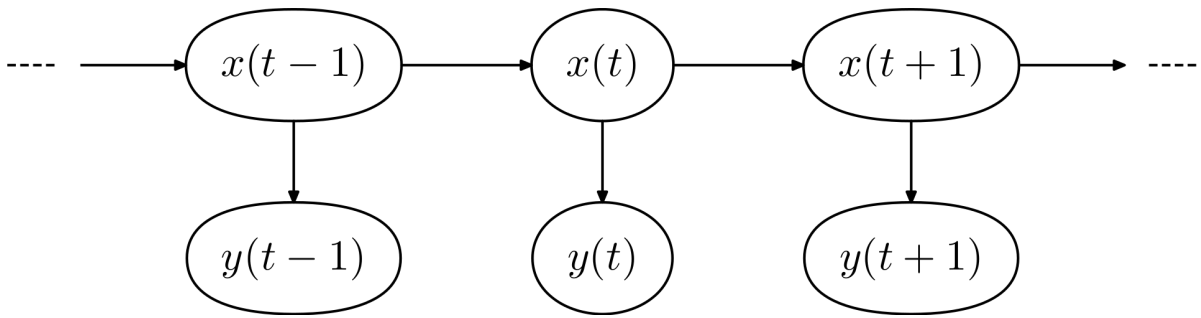
Characterizing HMMs



- Number of states
- Number of observations
- Emission probabilities
  - $P(Y_i|X_i)$
- Transition probabilities
  - $P(X_i|X_{i-1})$
- (Hidden) States
- Observations (data)
- Probability distribution(s)
- (Prior) Initial probabilities of states

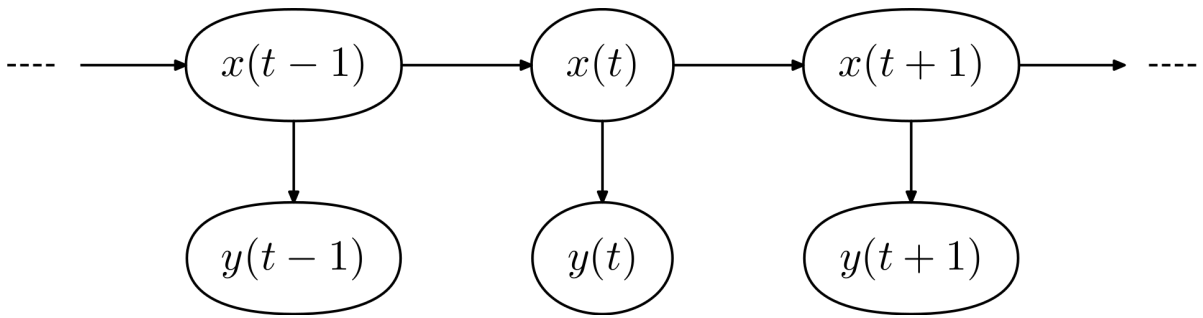
CAL POLY
SAN LUIS OBISPO

What Are We Interested In?



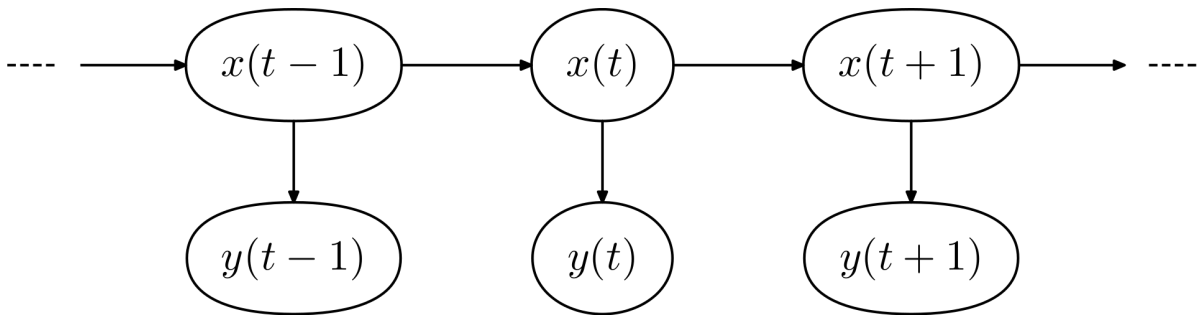
CAL POLY
SAN LUIS OBISPO

What Are We Interested In?



The States!!!

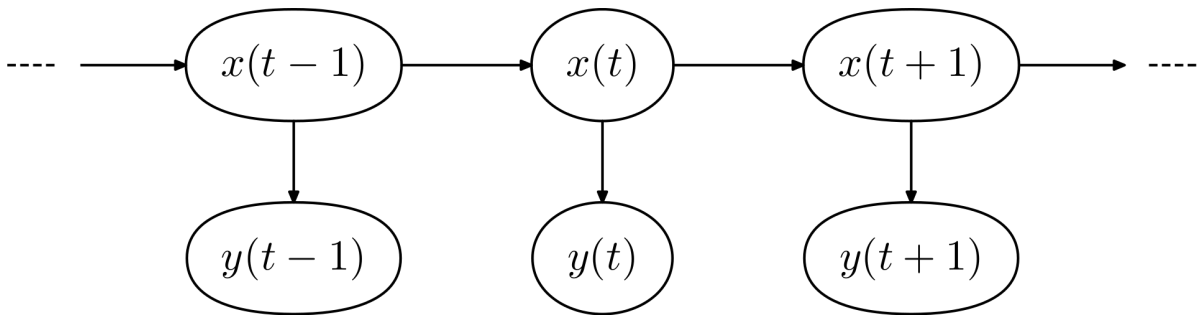
What Are We Interested In?



The States!!!

- How should we estimate the states?

What Are We Interested In?



The States!!!

- How should we estimate the states?
- Python Implementations
  - hmmlearn
  - seqlearn

CAL POLY
SAN LUIS OBISPO

HMM Inference

- Two things we might be interested in:

HMM Inference

- Two things we might be interested in:
  - Most likely sequence of hidden states (maximum a posteriori estimator)

$$\hat{X} = \underset{X}{\operatorname{argmax}}\, P(X|Y)$$

CAL POLY
SAN LUIS OBISPO

HMM Inference

- Two things we might be interested in:
  - Most likely sequence of hidden states (maximum a posteriori estimator)

$$\hat{X} = \underset{X}{\operatorname{argmax}}\, P(X|Y)$$

  - Centroid estimator (unconstrained)

$$\tilde{X}_i = \underset{X_i \in S}{\operatorname{argmax}}\, P(X_i|Y)$$

CAL POLY
SAN LUIS OBISPO

Maximum a Posteriori Estimator

- Small example…
- Viterbi Algorithm!

CAL POLY
SAN LUIS OBISPO

Centroid Estimator

- Continuing our small example….

Questions?

Questions?

CAL POLY
SAN LUIS OBISPO