Markov Models

Machine Learning
CSx824/ECEx242
Bert Huang
Virginia Tech

Last Time

- Probabilistic graphical models
- Bayesian networks
- Naive Bayes and Logistic Regression as Bayes nets
- Inference in Bayes nets

John Trumbull's *Declaration of Independence* in 1776

Independence

independent & identically distributed (i.i.d.)

full joint distributions

amount of dependence

cheap, easy, embarrassingly parallel

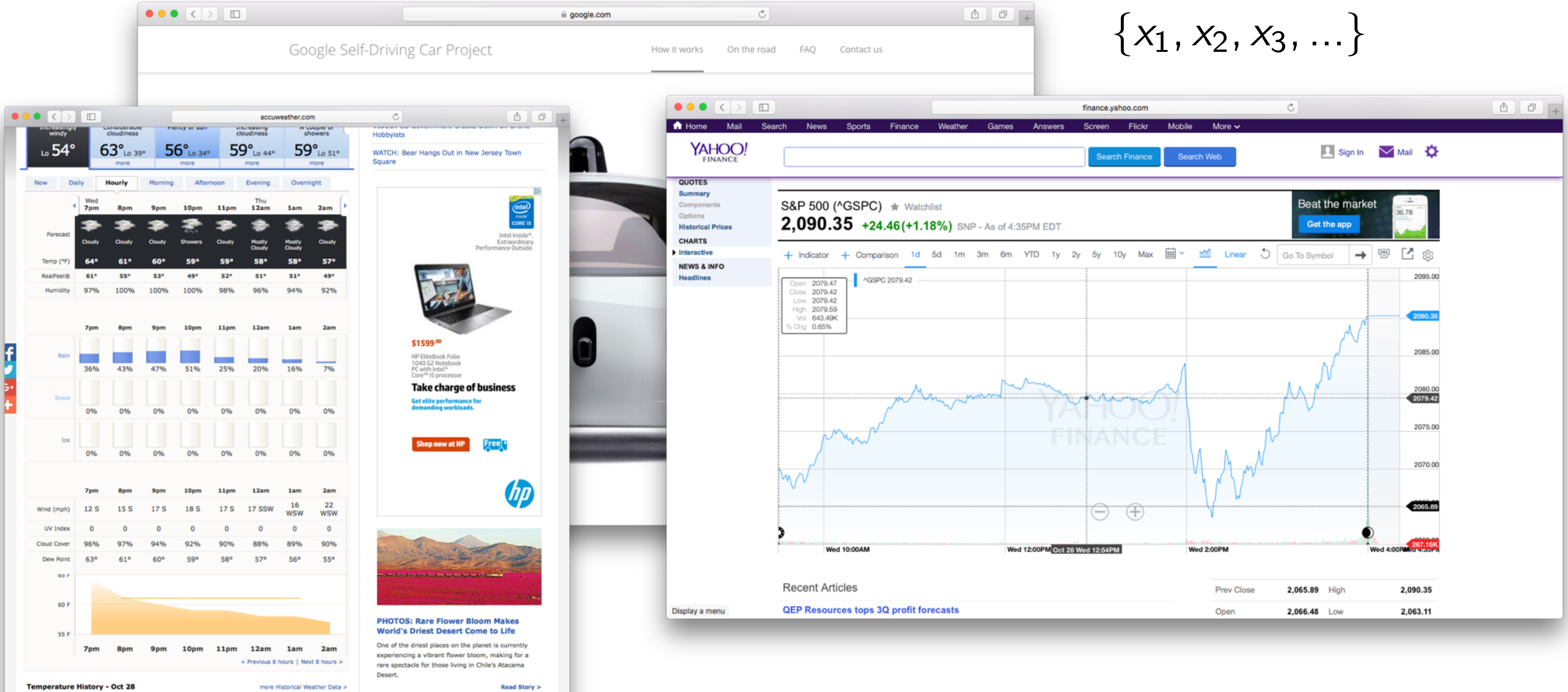
super expensive

Outline

- Time series
- Markov models
- Variable elimination in Markov models
- Forward message-passing inference

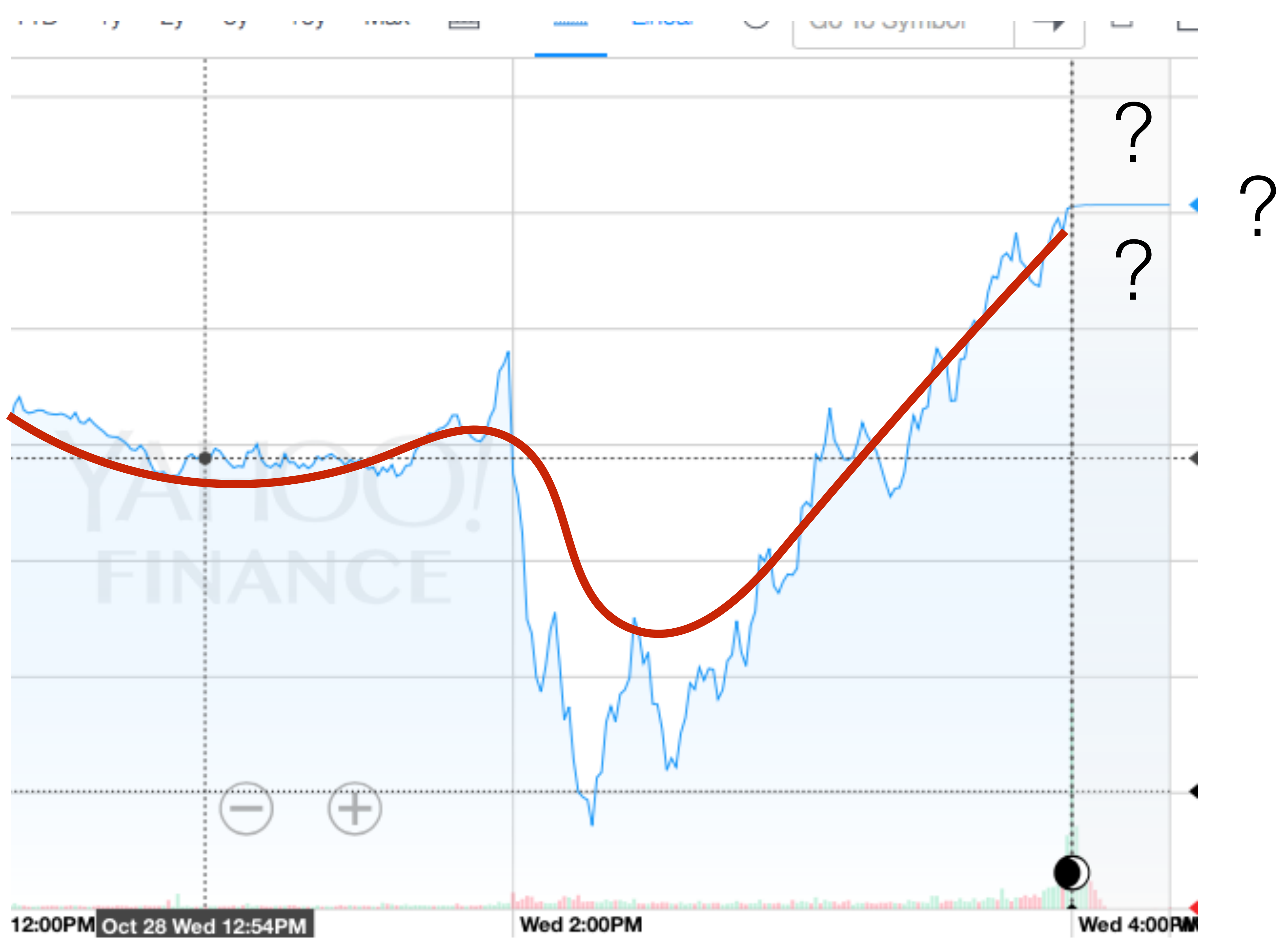
Time Series

$$\{x_1, x_2, x_3, \ldots\}$$



Time Series

- Goals:
  - Prediction
  - Filtering, smoothing



Markov Models

Markov assumption: the past is independent of the future given the present

$$p(x_i, x_k | x_j) = p(x_i | x_j) p(x_k | x_j) \qquad i < j < k$$

$$p(x_1, \dots, x_T) = p(x_1) \prod_{t=1}^{T-1} p(x_{t+1} | x_t)$$

usually parameterized with function independent of ***t***

Variable Elimination

$$p(x_1, x_2, x_3, x_4) = p(x_1)p(x_2|x_1)p(x_3|x_2)p(x_4|x_3)$$

$p(x_4)$?

$$p(x_4) = \sum_{x_1,x_2,x_3} p(x_1)p(x_2|x_1)p(x_3|x_2)p(x_4|x_3)$$

$$p(x_2) = \alpha_2(x_2) = \sum_{x_1} p(x_1)p(x_2|x_1)$$

$$p(x_4) = \sum_{x_2,x_3} \alpha_2(x_2)p(x_3|x_2)p(x_4|x_3)$$

$$p(x_3) = \alpha_3(x_3) = \sum_{x_2} \alpha_2(x_2)p(x_3|x_2)$$

$$p(x_4) = \sum_{x_3} \alpha_3(x_3)p(x_4|x_3)$$

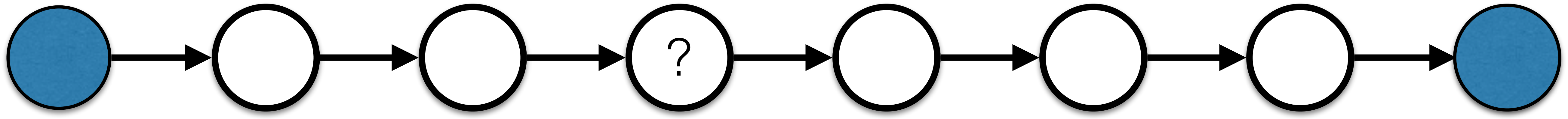
Forward Message Passing

$$p(X) = p(x_1) \prod_{t=1}^{T-1} p(x_{t+1}|x_t)$$

for ***t*** from 1 to (T-1):

$$p(x_{t+1}) = \sum_{x_t} p(x_t) p(x_{t+1}|x_t)$$

What about if we observe evidence?



$p(x_4|x_1 = 1)$ $p(x_4|x_1 = 1, x_8 = 0)$

Summary

- Time series
- Markov models
- Variable elimination in Markov models
- Forward message-passing inference