Announcements

- Final exam
  - Aug. 2, 6:20pm-8pm
  - Closed-book, 2 A4 cheat-sheets with your own writing
  - No electronic device, except basic calculator
- Still some questions?
  - Piazza
  - RC class on Thursday, 2-4pm
- Course evaluation

Advice

- Read carefully the problem description
  - Justify when needed
- Problems are independent
- Write clearly

Ve492: Introduction to Artificial Intelligence
Final Review

Paul Weng

UM-SJTU Joint Institute

Slides adapted from http://ai.berkeley.edu, AIMA, UM, CMU

Content

- Probability review
- Probabilistic reasoning
  - Bayes nets
  - Markov models and HMMs
- Machine learning
  - Naive Bayes
  - Perceptron
  - Neural networks
- Logic-based approaches
  - Propositional logic
  - First-order logic
  - Classical planning

Probability

- For each of the following statements, either prove it is true or give a counterexample.
  - If $P(a \mid b, c) = P(b \mid a, c)$, then $P(a \mid c) = P(b \mid c)$
  - If $P(a \mid b, c) = P(a)$, then $P(b \mid c) = P(b)$
  - If $P(a \mid b) = P(a)$, then $P(a \mid b, c) = P(a \mid c)$

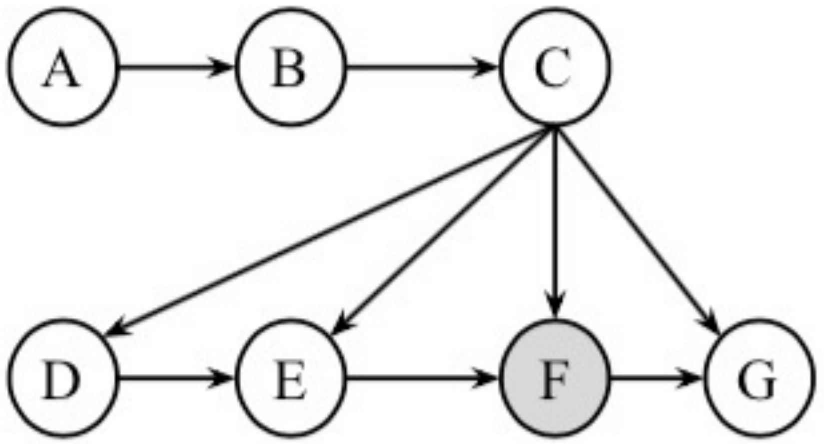
Bayes Rule

Consider two medical tests, A and B, for a virus. Test A is 95% effective at recognizing the virus when it is present, but has a 10% false positive rate (indicating that the virus is present, when it is not). Test B is 90% effective at recognizing the virus, but has a 5% false positive rate. The two tests use independent methods of identifying the virus. The virus is carried by 1% of all people. Say that a person is tested for the virus using only one of the tests, and that test comes back positive for carrying the virus.

- Which test returning positive is more indicative of someone really carrying the virus? Justify your answer mathematically.

Bayes' Net

- Write the joint distribution of the following Bayes' net

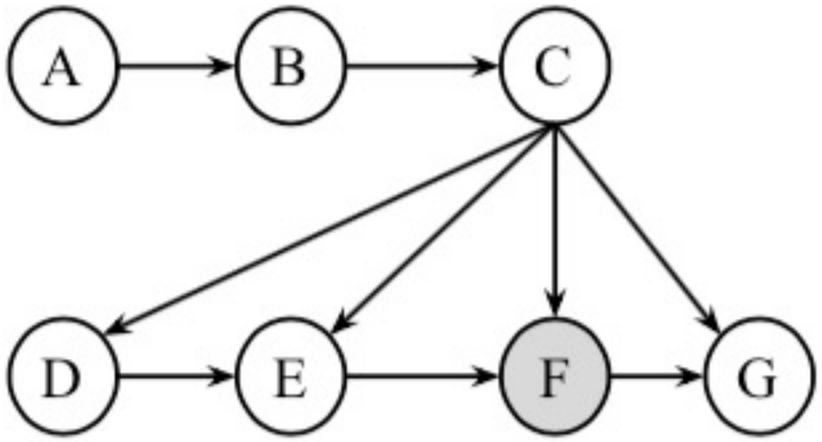


- How many values does the joint distribution have?
- How many parameters does the Bayes' net have?

D-Separation

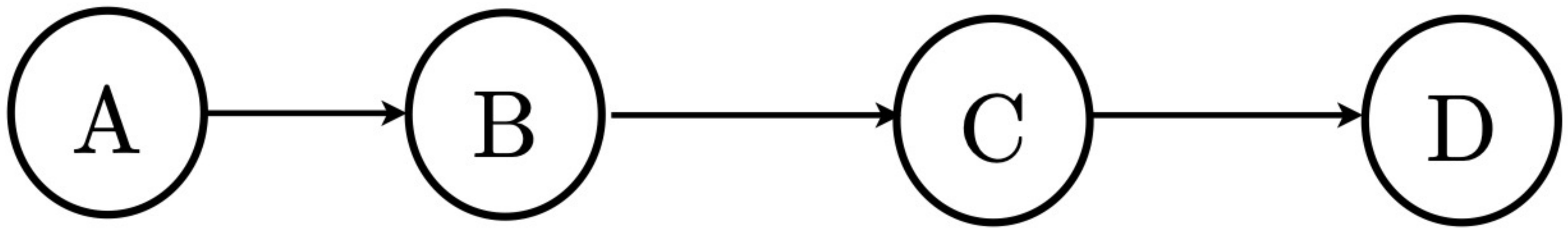
- Check Bayes applet

Inference



- Run Variable Elimination to compute $P(B, D| + f)$ with order A, C, E, G
- What is the size of the largest generated factor?
- Find the best ordering for Variable Elimination
- What is the cutset for this graph?

Sampling



$P(A)$	
$-a$	3/4
$+a$	1/4

$P(B\|A)$		
$-a$	$-b$	2/3
$-a$	$+b$	1/3
$+a$	$-b$	4/5
$+a$	$+b$	1/5

$P(C\|B)$		
$-b$	$-c$	1/4
$-b$	$+c$	3/4
$+b$	$-c$	1/2
$+b$	$+c$	1/2

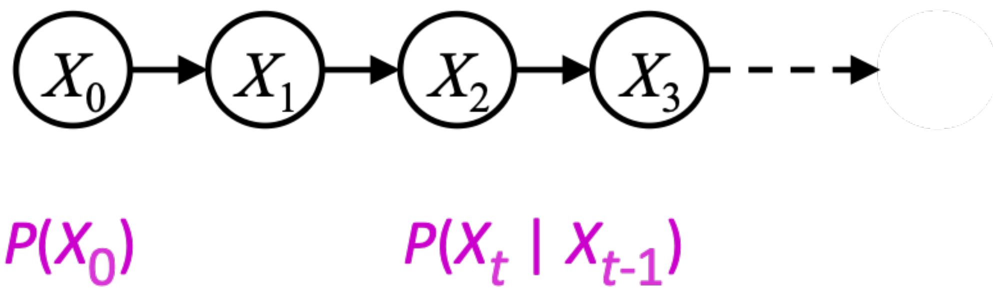
$P(D\|C)$		
$-c$	$-d$	1/8
$-c$	$+d$	7/8
$+c$	$-d$	5/6
$+c$	$+d$	1/6

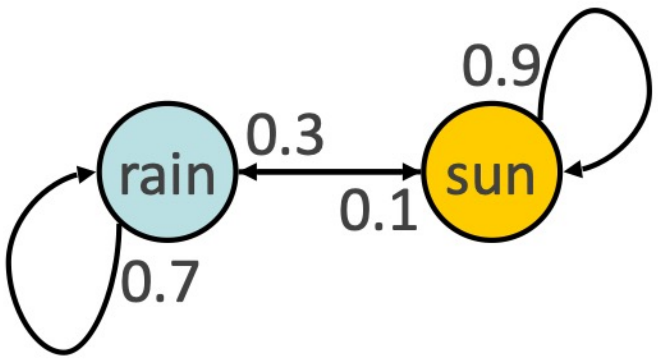
Samples

$+a$	$+b$	$-c$	$-d$
$+a$	$-b$	$+c$	$-d$
$-a$	$+b$	$+c$	$-d$
$-a$	$-b$	$+c$	$-d$
$+a$	$-b$	$-c$	$+d$
$+a$	$+b$	$+c$	$-d$
$-a$	$+b$	$-c$	$+d$
$-a$	$-b$	$+c$	$-d$

- Estimate $P(+c| +a, -d)$ via rejection sampling
- Estimate $P(-a| +b, -d)$ via likelihood weighting

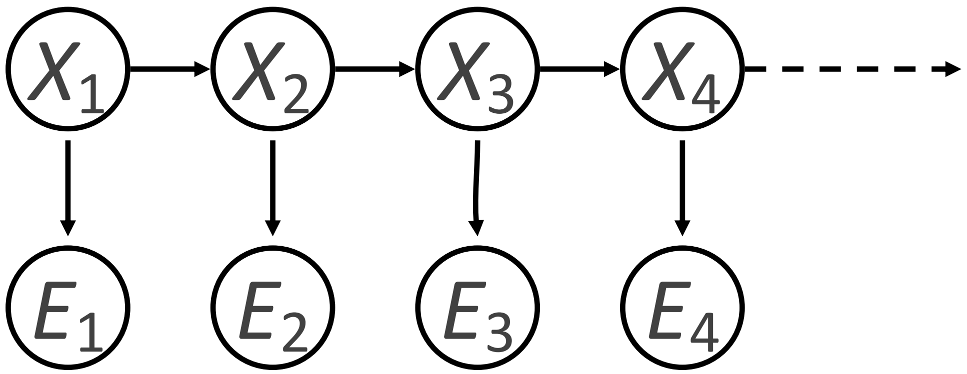
Markov Chain

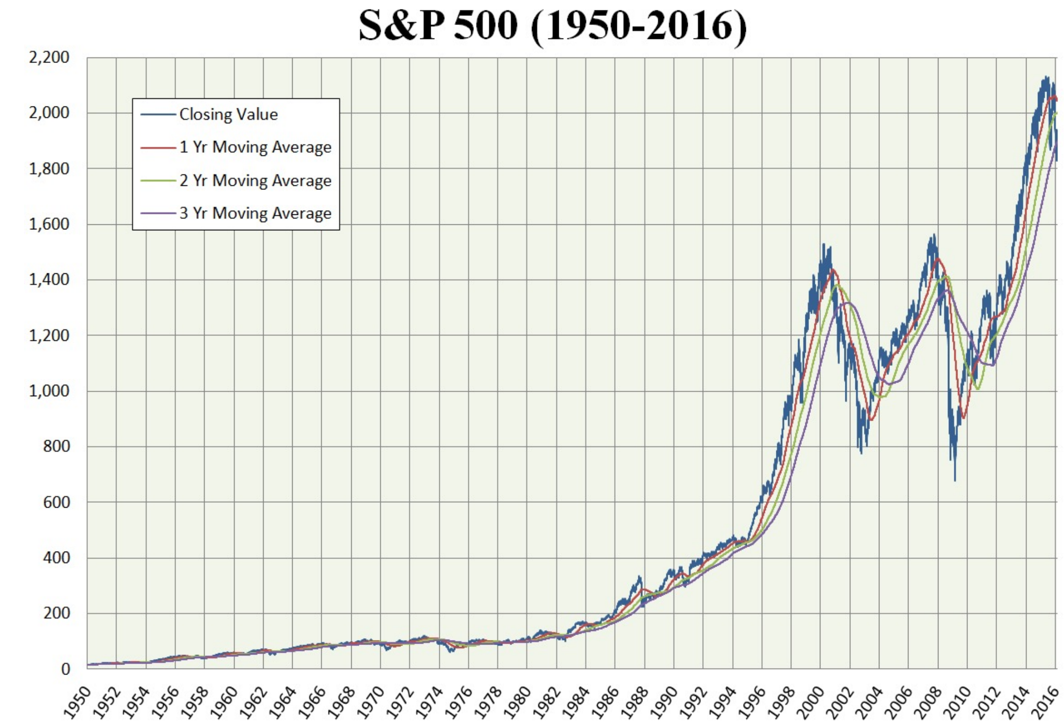




- What is the probability of $P(X_t)$?
- What is the stationary distribution for the weather example?

Hidden Markov Model

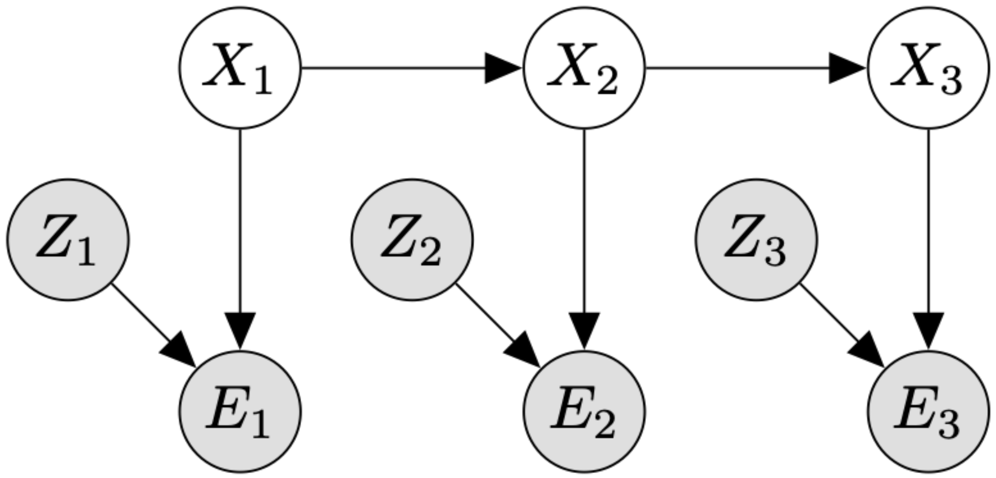




Wikipedia

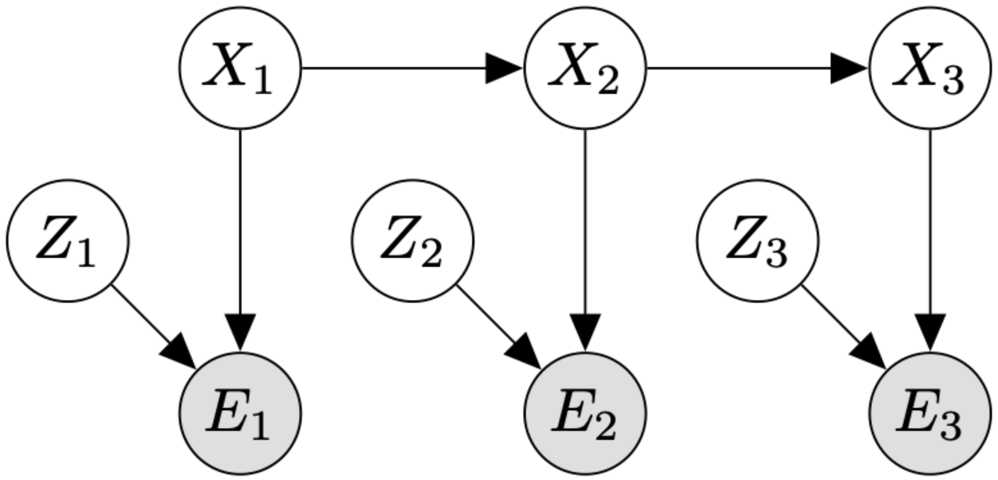
- Financial investment
  - X = market condition: bull, bear
  - E = price evolution of some index: up, down
- Label past data into bull vs bear
- Use historical data to estimate transition/emission probabilities

Hidden Markov Model



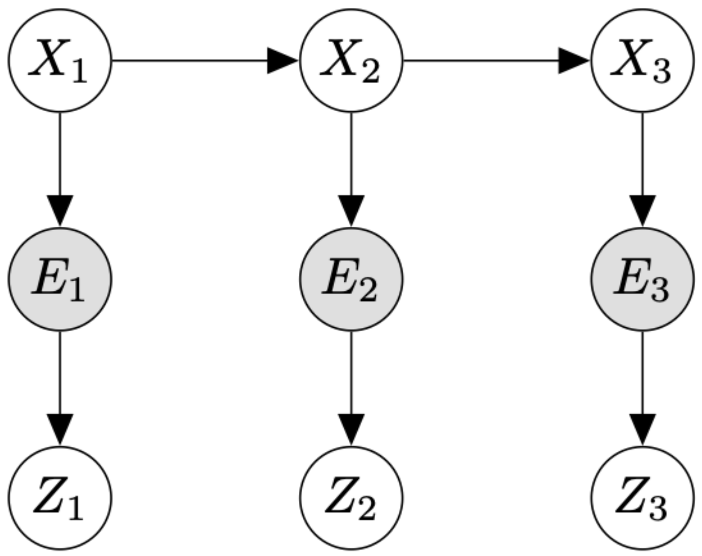
- Adapt the forward algorithm to this variant of HMM
  - Predict step
  - Update

Hidden Markov Model



- Adapt the forward algorithm to this variant of HMM
  - Predict step
  - Update

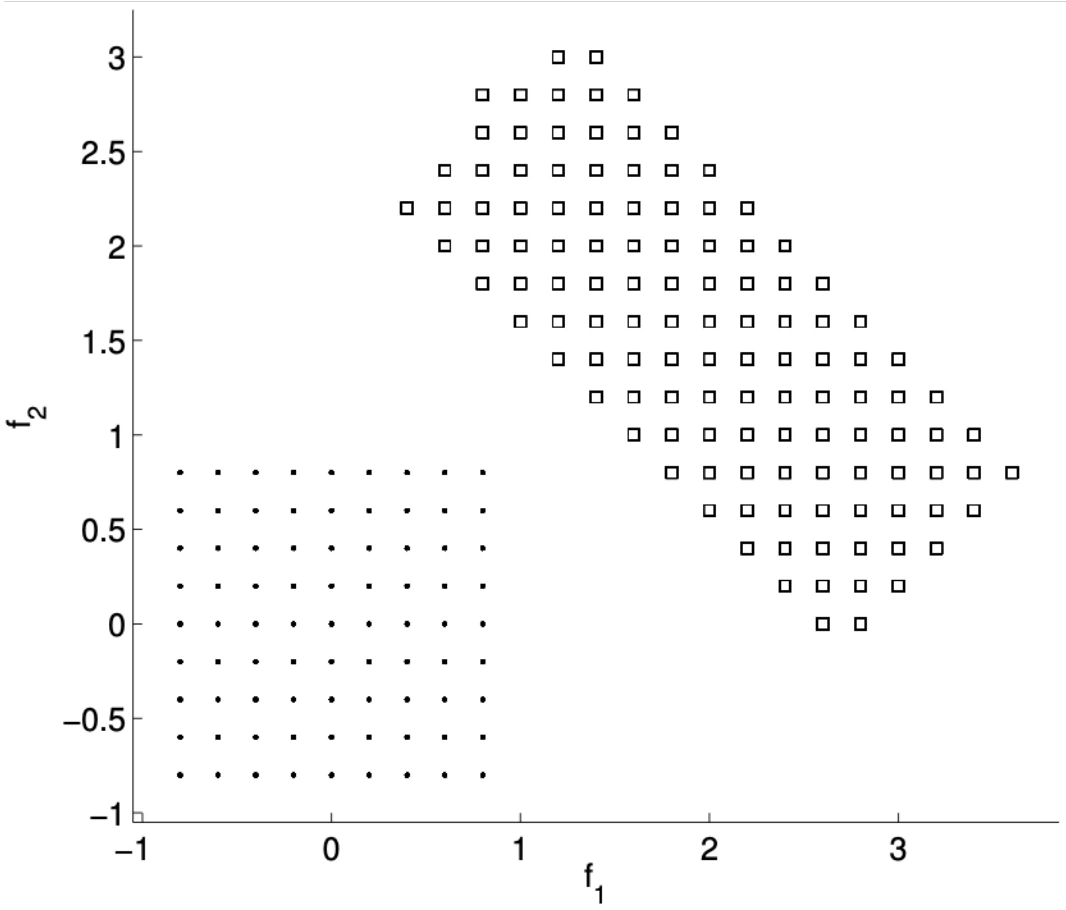
Hidden Markov Model

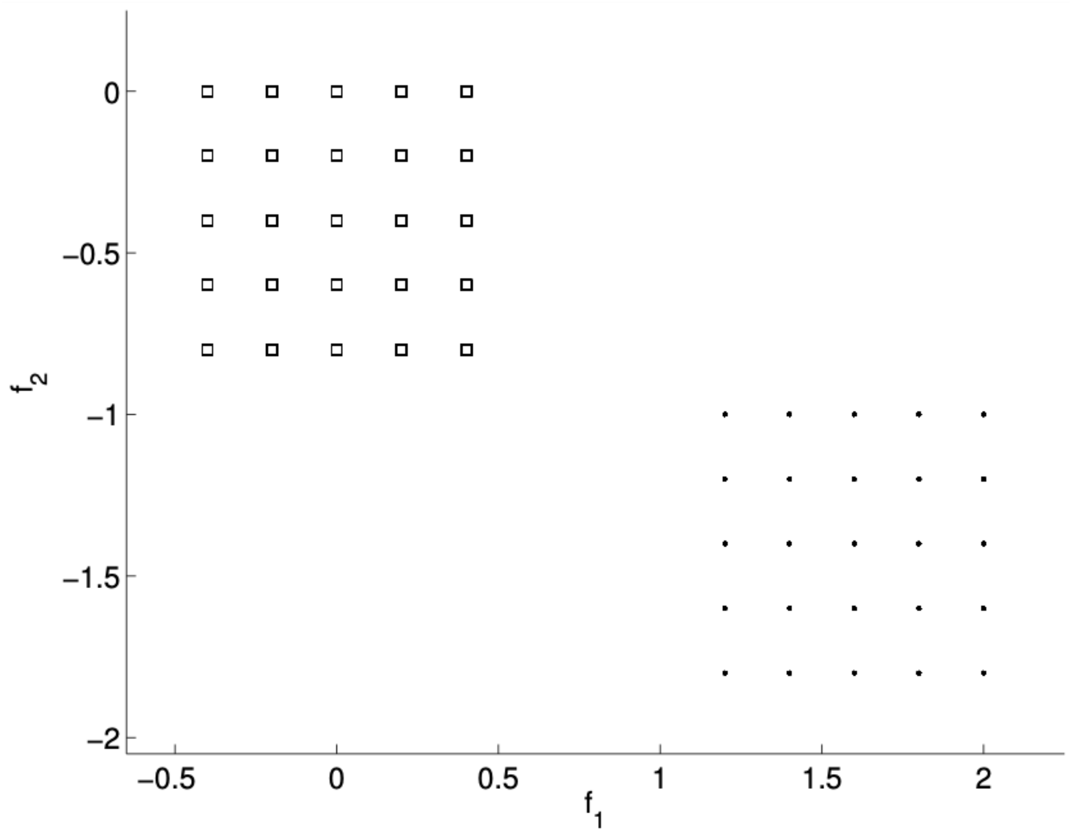


- Adapt the forward algorithm to this variant of HMM
    - Predict step
    - Update

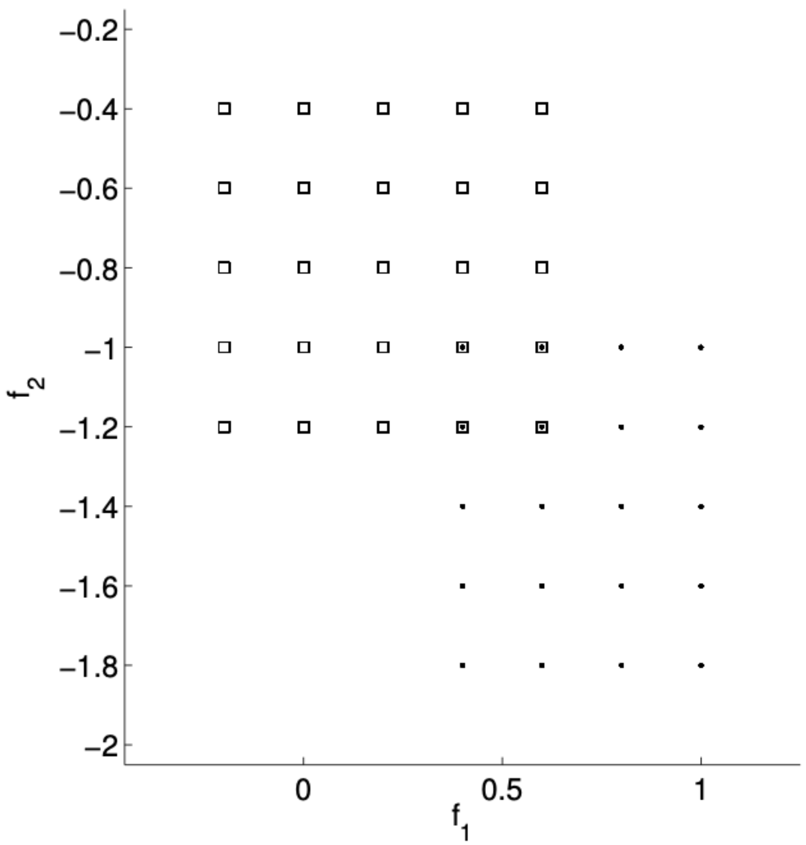
Naïve Bayes

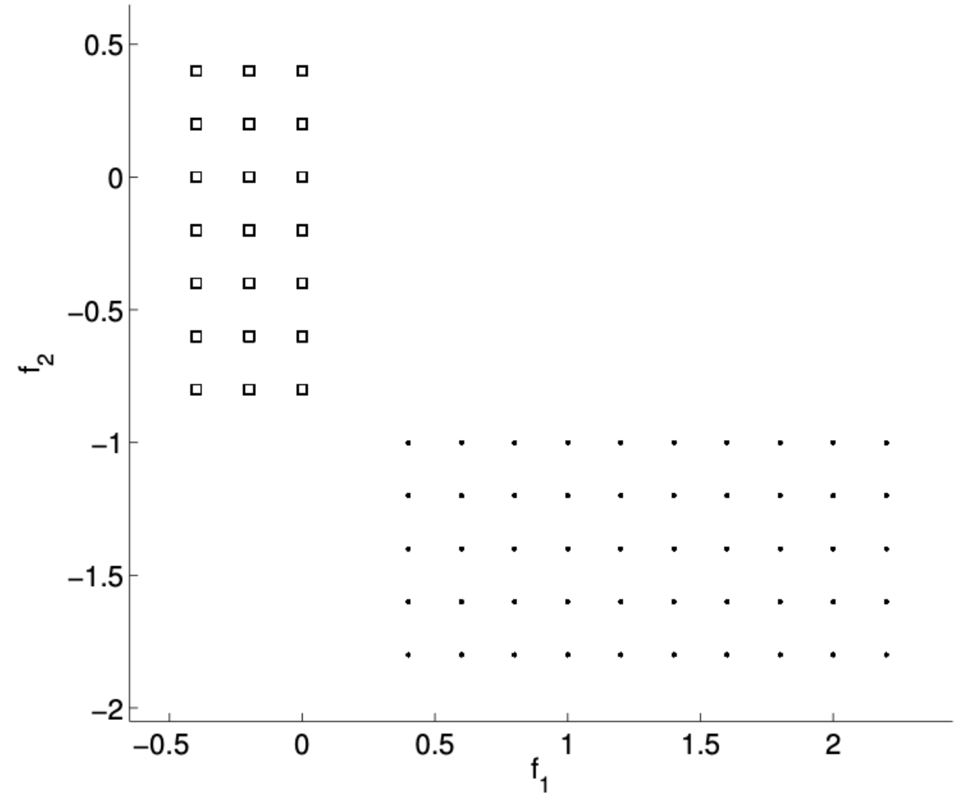
- Which of the following binary classification problems satisfy the assumption made in Naïve Bayes?

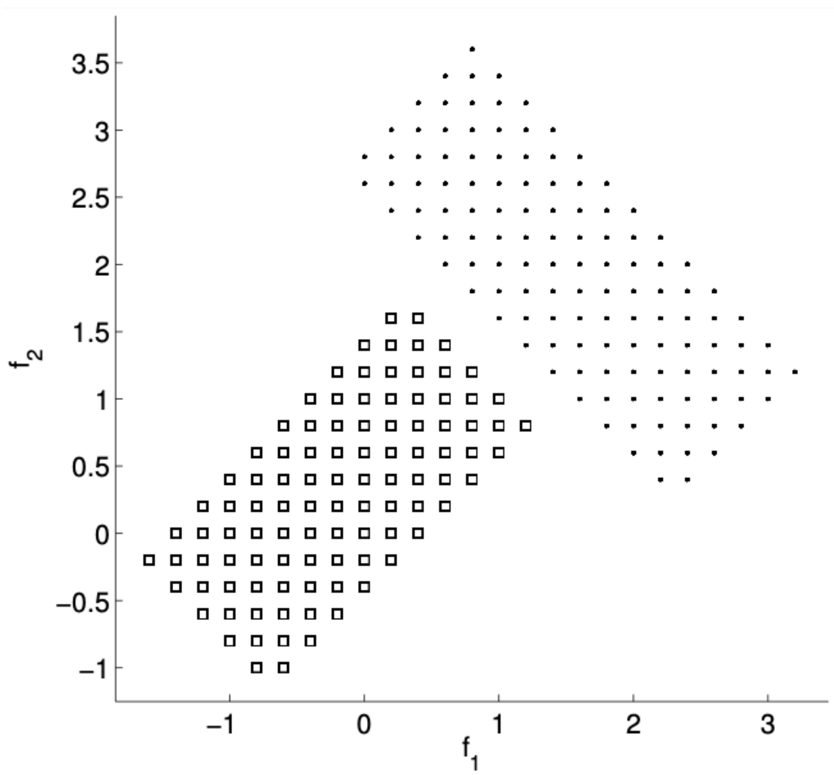


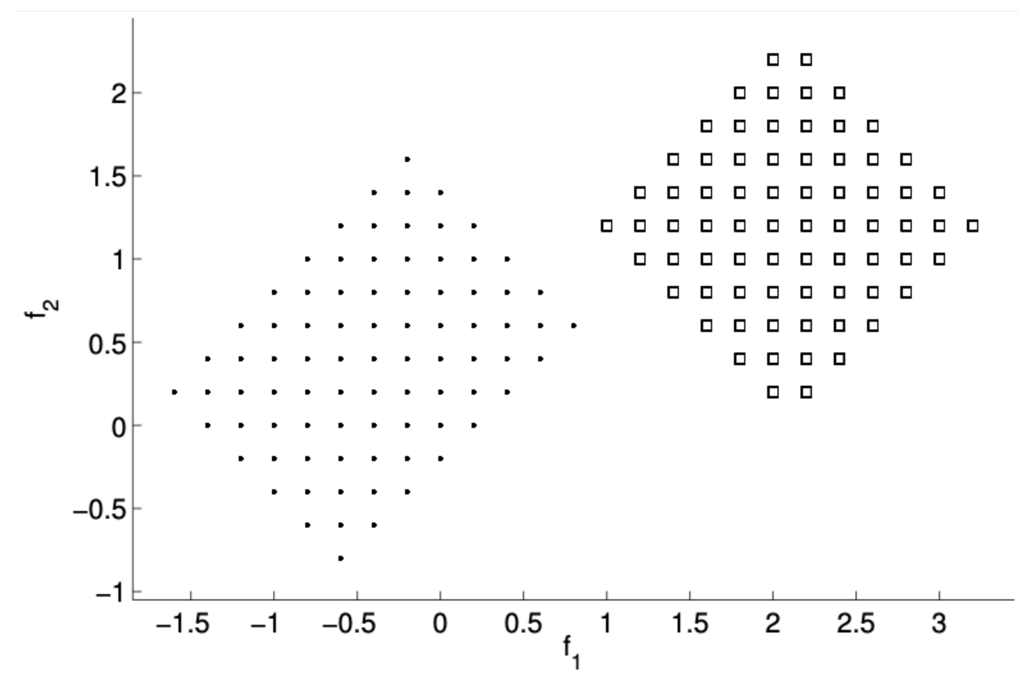


Naïve Bayes ctd.









Discriminative Learning

For a binary classification problem, we choose the following model $\mathbb{P}(y = +1|x) = \Phi(w \cdot x)$ where Φ is the CDF of a standard normal distribution.

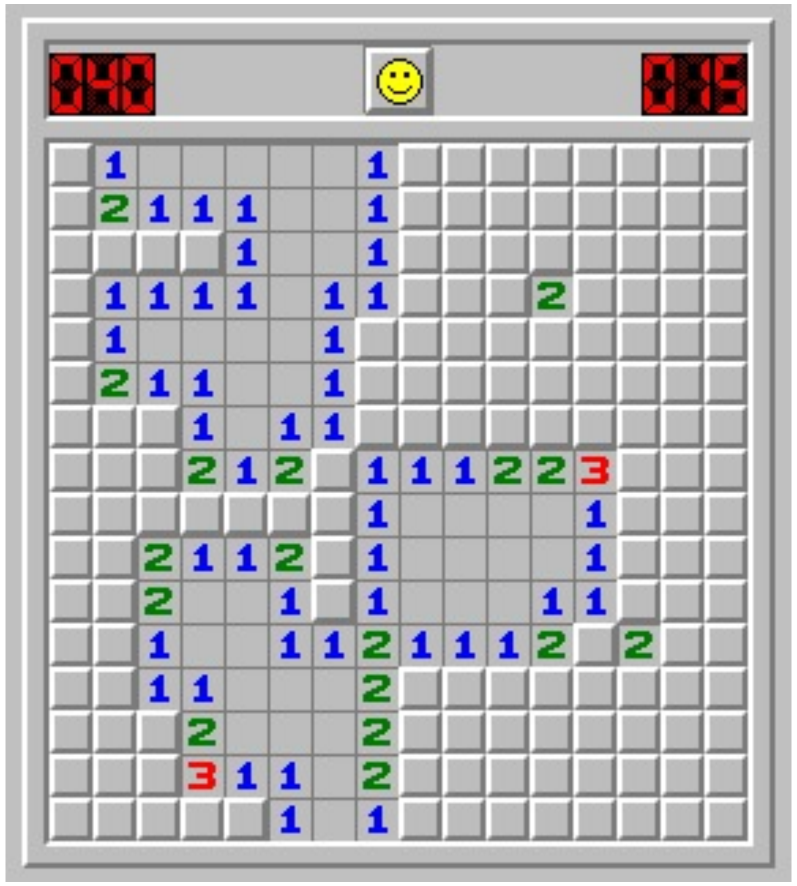
- What is the decision boundary?
- Formulate the optimization problem to be solved to find w
- Formulate the stochastic gradient method to solve that problem

Propositional Logic

- Which of the following are correct?
  - $False \models True$
  - $True \models False$
  - $(A \wedge B) \models (A \Leftrightarrow B)$
  - $(A \vee B) \wedge (\neg C \vee \neg D \vee E) \models (A \vee B)$
  - $(A \vee B) \wedge \neg(A \Rightarrow B)$ is satisfiable
  - $(A \Leftrightarrow B) \wedge (\neg A \vee B)$ is satisfiable
  - $(A \Leftrightarrow B) \Leftrightarrow C$ has the same number of models as $(A \Leftrightarrow B)$ for any fixed set of proposition symbols that includes A, B, C

Application: Propositional Logic

- Minesweeper: Let $X_{i,j}$ be true iff square $[i, j]$ contains a mine.
  - Write down the assertion that exactly two mines are adjacent to [1,1] as a sentence involving some logical combination of $X_{i,j}$ propositions.
  - Generalize your assertion by explaining how to construct a CNF sentence asserting that k of n neighbors contain mines
  - Explain precisely how an agent can use DPLL to prove that a given square does (or does not) contain a mine.



First-Order Logic

- For each of the following sentences in English, decide if the accompanying first-order logic sentence is a good translation. If not, explain why not and correct it.
  - No two people have the same social security number.

    $\neg\exists$ x, y, n Person(x) $\wedge$ Person(y) $\Rightarrow$ [HasSS#(x, n) $\wedge$ HasSS#(y, n)]

  - John's social security number is the same as Mary's.

    $\exists$ n HasSS#(John, n) $\wedge$ HasSS#(Mary, n)

  - Everyone's social security number has nine digits.

    $\forall$ x, n Person(x) $\Rightarrow$ [HasSS#(x, n) $\wedge$ Digits(n, 9)]

Classic Planning

A monkey is in a room with some bananas hanging out of reach from the ceiling. A box is available that will enable the monkey to reach the bananas if he climbs on it. Initially, the monkey is at *A*, the bananas are at *B*, and the box is at *C*. The monkey and the box have height *Low*, but if the monkey climbs onto the box he will have height *High*, the same as the bananas. The actions available to the monkey include *EatBananas* if the monkey and the bananas are at the same location and height, *Go* from one place to another, *Push* an object from one place to another, and *ClimbUp* onto or *ClimbDown* from an object.

- Write down the initial state description
- Write down the STRIPS definitions of the five actions.