Q1 Score: **3** Total Score: **65.0**

88D8E05D-3637-48AA-93F5-B15B749DF197

SPCSCI561-examFinalg

#27 2 of 19

1. Multiple Choice Qeustions (5 points)

D 1. This is an advantage of symbolic methods in AI over statistical approaches:
   a. Good for inherently uncertain or approximate domains
   b. More robust to unanticipated situations
   c. Can work at very large scale
   d. Can explain reasoning
   e. All of the above
   f. None of the above

A 2. This is a disadvantage of statistical methods in AI when compared with symbolic approaches:
   a. Hard to know what to do if no data
   b. Good for only small domains, small datasets
   c. Fragile
   d. Machine learning more difficult
   e. All of the above
   f. None of the above

F 3. Which is NOT an example of an instance of cyberhumanism:
   a. Skype
   b. Siri
   c. Fitbit
   d. Google Glass
   e. All of the above
   f. None of the above

D 4. Which is NOT a level of analysis in natural language processing?
   a. Morphology
   b. Phonology
   c. Pragmatics
   d. Physiology
   e. All of the above
   f. None of the above

E 5. Why are operational models of emotion needed to make virtual humans work?
   a. Reasoning about verbal and non-verbal communication and perception
   b. Shapes decision-making
   c. Emotions expressed through behavior reveals mental state and intentions
   d. Emotions can change behavior of observers
   e. All of the above
   f. None of the above

Q2 Score: **5**

2. True/False Questions (10 points)

T **F** 1. [True/False] Cyberhumanism: Social movement that looks forward to future in which humans are merged with – or replaced by robotic and biotech inventions

T **F** 2. [True/False] Singularity, Transhumanism and virtual humans are all approaches to designing information technology that seeks to work better with humans through better understanding of the interests, needs, capabilities and welfare of humans.

F **T** 3. [True/False] AI planning and machine vision techniques are used in emotion appraisal systems

F **T** 4. [True/False] An agent evoking social effects can produce negative outcomes

F **F** 6. [True/False] Naive Bayes classification for text categorization on shorter texts (utterances, tweets), usually has much higher performance than on longer documents.

T **T** 7. [True/False] Understanding emotions may be important for machines that do not deal with people.

T **F** 8. [True/False] Machine Translation and Dialogue Systems are two application areas for natural language processing that do not use statistical AI techniques.

F **F** 9. [True/False] Training data of text labeled with polarity for Sentiment analysis is hard to find.

F **T** 10. [True/False] Parsing for Information Extraction and Information Retrieval uses greedy search techniques for accuracy and speed.

Q3 Not graded

3. Planning (15 points)

Consider the following STRIPS action schema for planning an end of semester party:

LookupExamDate(subject)
  Pre: Enrolled(subject)
  Post: HaveExamDate(subject)

Schedule(party)
  Pre: HaveExamDate(AI), HaveExamDate(Algorithms)
  Post: HaveDate(party)

InviteGuests()
  Pre: HaveDate(Party)
  Post: Invited(Guests)

GetAttendanceEstimate(party)
  Pre: Invited(Guests)
  Post: HaveAttendance(party)

ShopPartyRefreshments()
  Pre: Prepared(AI), Prepared(Algorithms), HaveAttendance(party)
  Post: Have(PartySupplies)

CleanHouse(house)
  Pre: Prepared(AI), Prepared(Algorithms)
  Post: Clean(house)

Study(subject)
  Pre:
  Post: Prepared(subject)

TakeExam(subject)
  Pre: Prepared(subject), HaveExamDate(subject)
  Post: CompletedExam(subject)

HaveParty(party)
  Pre: Invited(Guests), CompletedExam(AI), CompletedExam(Algorithms),
    Have(PartySupplies), Clean(House)
  Post: Enjoyed(party)

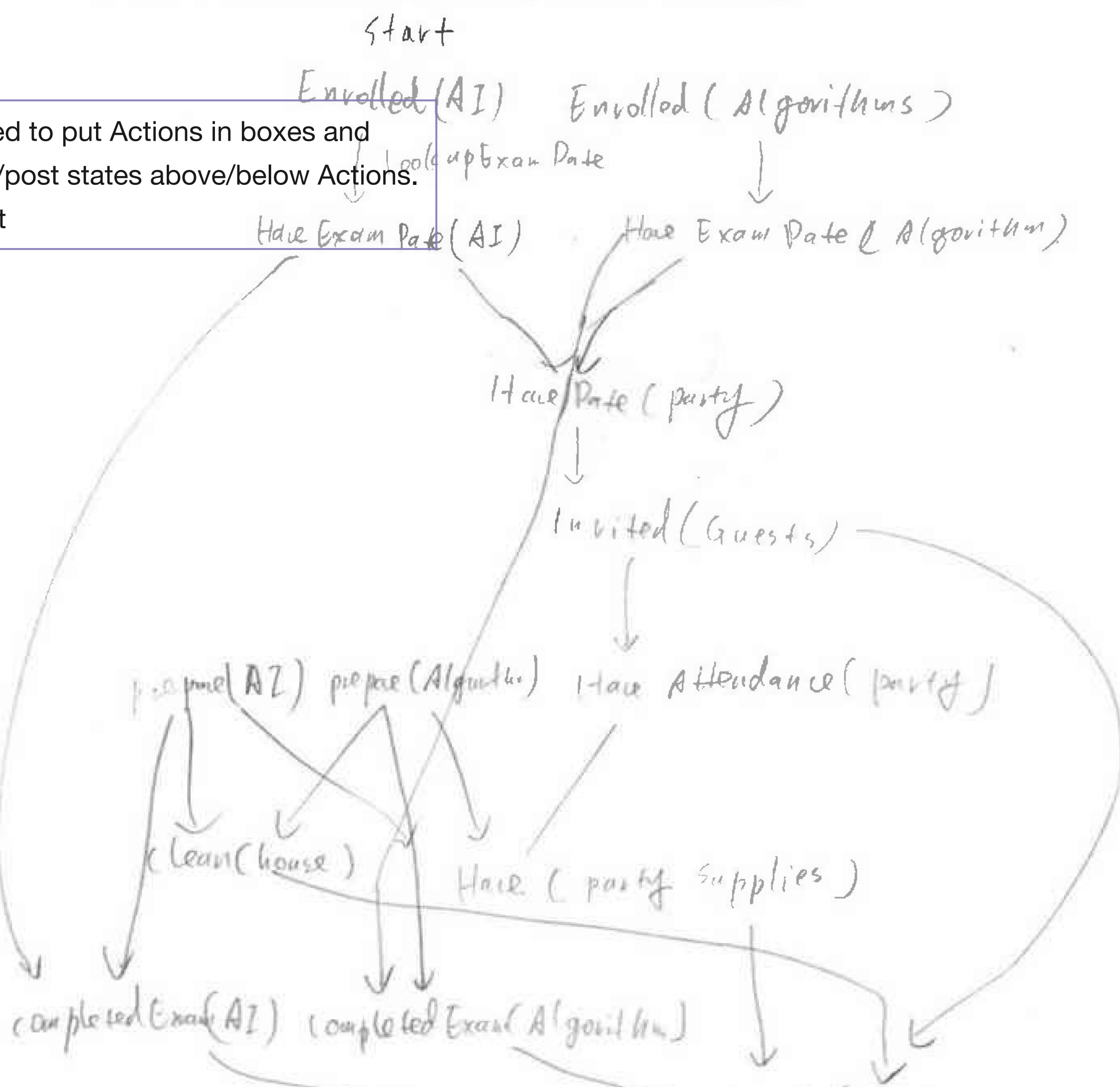
Enroll
Have Exam Date AI, Al.
Have Date
Invited
Study
prepared(AI, Al) Have Attendance
Clean
Have
Complete Exam AI, Al
Enjoyed

Q4 Score: **13**

Given the following:

Start: Enrolled(AI); Enrolled(Algorithms)
End: Enjoyed(EndOfTermParty)

Show the partial order plan for planning the end of semester. Your answer should be a complete plan in the form of a graph showing actions from the Start state to the Finish state using the actions above. Your answer should reflect the ordering constraints required by the operators, but should also take advantage of partial orderings, allowing the plan executor as much choice as possible when the action ordering is not constrained.
Note: You have an extra page to show your work if necessary.



Need to put Actions in boxes and pre/post states above/below Actions.
-2pt

Q6 Score: 10

$P(a\,b) = P(a|b)\,P(b) = P(b|a)\,P(b)$

$P(a|b) = P(b|a)\,P(a)/P(b)$

$P(a|b) = P(a\,b)/P(b)$

$P(a|b)\,P(b) = P(b|a)\,P(a)$

620C6A0B-8905-465A-BAE2-DB405F289833
SPCSCI561-examFinalg
#27 7 of 19

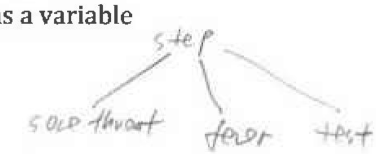
4. Bayesian Probabilities (22 points)

You wake up one morning with a bad sore throat. You might have strep throat, which is caused by a streptococcal infection, or it might just be a cold. It's good to know what the cause is because the strep infection can be treated with antibiotics, on the other hand doctors are reluctant to give everyone with a sore throat antibiotics because it leads to over-use. Looking on the web, you find that for the population of people with sore throat, the probably they have strep is 15%:

~~P strep | sore th~~

That is, P(strep | sore throat) = 0.15.

Because in this problem we are going to be focusing exclusively on just the population of people with sore throats, we can remove "sore throat" as a variable and consider just the probability of strep for that population, that is,



P(strep) = 0.15

1. You also learn that there are other symptoms that can be indicative of strep:

P(fever | strep) = 0.6
P(fever | ¬strep) = 0.3

$\frac{0.6 \quad 0.05}{30}$ $\quad 0.03 \quad$ $\frac{0.3 \quad 0.9}{0.27}$ $\quad \frac{0.03 \quad 0.15}{0.3} = 0.015$

a) What is the probability that you have strep if you have a fever as well, that is: P(strep | fever) ? (2 points)

$P(strep|fever) = \frac{P\,st(fever\ strep)}{P(fever)} = \frac{P(fever|strep)\,P(strep)}{P(fever|strep) + P(fever|\neg strep)} = \frac{0.6\ 0.15}{0.6 + 0.3}$

b) What if you don't have a fever? P(strep | ¬fever)? (2 points)

$P(strep|\neg fever) = \frac{P(\neg fever|strep)\,P(strep)}{P(\neg fever|strep) + P(\neg fever|\neg strep)} = \frac{0.4\ 0.15}{0.4\ 0.7} = \frac{3}{14}$

$= \frac{0.09}{0.9} = \frac{1}{10}$

2. A new diagnostic test has been developed for identifying the strep organism.
P(test | strep) = 0.95
P(test | ¬strep) = 0.10

Assuming the test and fever are conditionally independent given strep, what is:

a) P(strep | fever, test) ? (3 points)

$P(strep|fever, test) = \frac{P(strep, fever, test)}{P(fever, test)} = \frac{P(fever|strep)\,P(test|strep)\,P(strep)}{P(fever|strep)\,P(test|strep) + P(fever|\neg strep)\,P(test|\neg strep)}$

$= \frac{0.6 \cdot 0.95 \cdot 0.15}{0.6 \cdot 0.95 + 0.3 \cdot 0.1} = \frac{0.0855}{0.6}$

b) P(strep | fever, ¬test) ? (3 points)

$P(strep|fever, \neg test) = \frac{P(fever|strep)\,P(\neg test|strep)\,P(strep)}{P(fever|strep)\,P(\neg test|strep) + P(fever|\neg strep)\,P(\neg test|\neg strep)}$

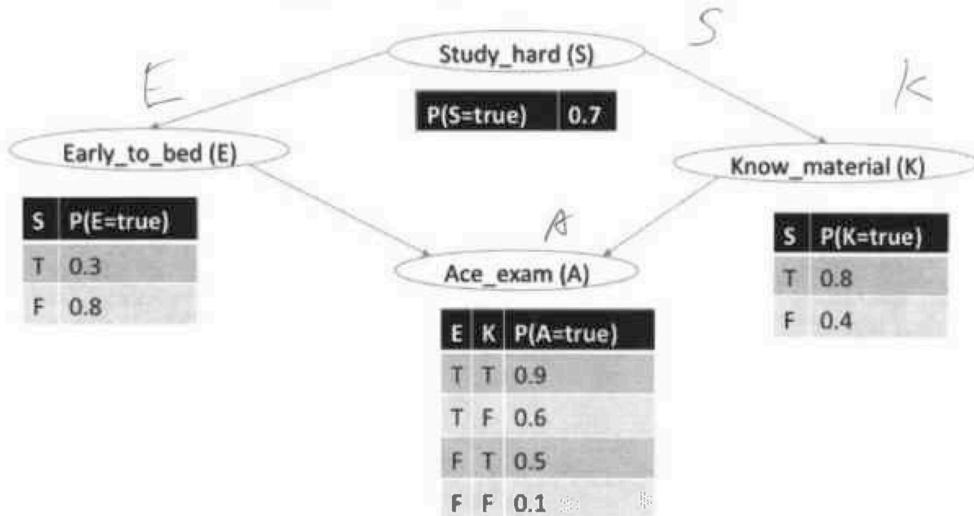
$= \frac{0.6\ 0.05\ 0.15}{0.6\ 0.05 + 0.3 \cdot 0.9} = 0.015$

Q8 Score: **7**

85AE8FD8-CFD5-45A9-8148-EED849BE793A
SPCSCI561-examFinalg
#27 9 of 19

Bayes Nets

3. Given the following Bayes net:
Note: You have an extra page to show your work if necessary.



P(S=true)	0.7

S	P(E=true)
T	0.3
F	0.8

S	P(K=true)
T	0.8
F	0.4

E	K	P(A=true)
T	T	0.9
T	F	0.6
F	T	0.5
F	F	0.1

a) What is P(S, E, K, ¬A)? (Show all work) (3 points)

$= P(S) \cdot P(E|S)\ P(K|S)\ P(\neg A|EK)$

$= 0.7 \cdot 0.3\ \ 0.8 \cdot\ 0.1$

$= 0.0168$

b) What is P(K | S, E)? (2 points)

$= \frac{P(K,S,E)}{P(S,E)} = \frac{P(K|S)\ P(E|S)\ P(S)}{P(E|S)\ P(S)} = \frac{0.8\ \ 0.3\ \ 0.7}{0.3\ \ 0.7} = 0.8$

c) What is P(S | A) ? (You do not need to work out the numeric computation on this part, just show the formulas and plug in the values from the tables that would give the result) (3 points)

$P(S|A) = \frac{P(S,A)}{P(A)} = \frac{P(S, A, K=k, E=e)}{P(A, S=s, E=e, K=k)}$

Q9 Score: **0**

C612E190-CC24-4496-99ED-741C1078C998
SPCSCI561-examFinalg
#27 10 of 19

Use this page to show results from preceding problem, if necessary.

$P(S, A, K=k, E=e)$

$= \sum_{K=k, E=e} \left(P(S) P(E|S) P(K|S) P(A|EK) \right)$

$K=K, \neg K \quad E=E, \neg E$

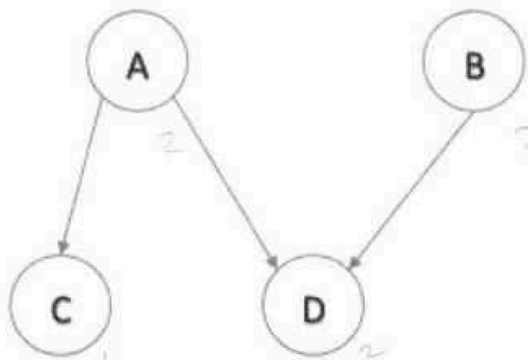
$P(A, S=s, E=e, K=k)$

$= \sum_{K=k\ E=e\ S=s} \left(P(S) P(E|S) P(K|S) P(A|EK) \right)$

$K=K, \neg K, \quad E=E, \neg E \quad S=S, \neg S.$

Q10 Score: **2**

4. In this Bayesian network, A, B, C and D are all binary random variables.



a) Considering all of the conditional probability tables (CPTs) that would be associated with this network, what is the minimum number of probability entries needed for A, B, C, and D? (2 points)

7.

b) Construct another Bayesian network for variables A, B, C and D consistent with the following: (2 points)

- A is independent of B
- C depends on A and B
- D depends on C
- D is conditionally independent of A and B given C

Q11 Score: **2**

$I\left(\frac{p}{p+n}, \frac{n}{p+n}\right) = -\frac{p}{p+n}\log_2\frac{p}{p+n} - \frac{n}{p+n}\log_2\frac{n}{p+n}$

AFAD3305-54D9-4E37-8DD0-6DE9B69545D3

SPCSCI561-examFinalg

#27 12 of 19

5. Decision Tree Learning (15 points)

1. You are given a decision problem in which there are 12 "True" examples and 4 "False" examples. What is the information *I* (also called entropy) conveyed by this distribution? Show the formula for calculating with numbers plugged in. You do not need to work out the numeric calculation. (3 points)

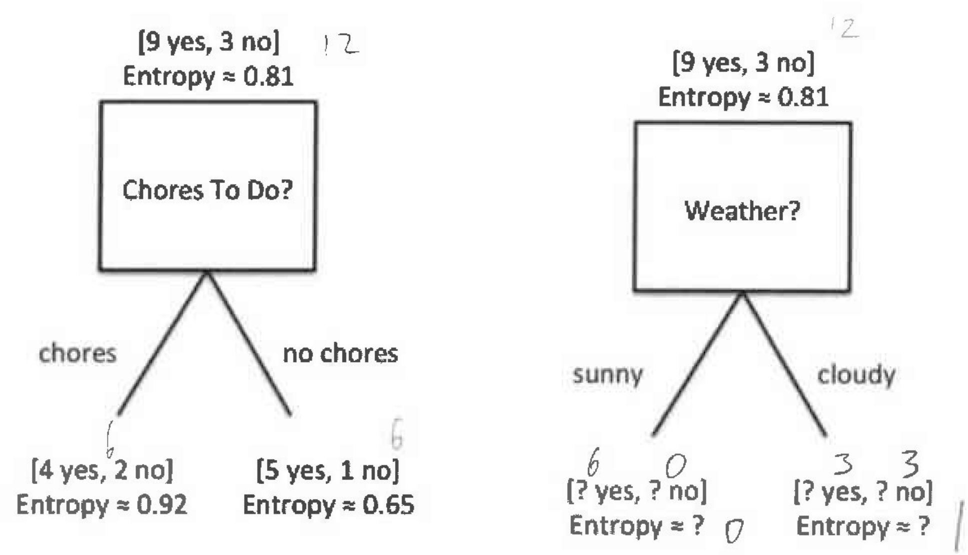
I(P) = $I\left(\frac{p}{p+n}, \frac{n}{p+n}\right) = -\frac{p}{p+n}\log_2\frac{p}{p+n} - \frac{n}{p+n}\log_2\frac{n}{p+n}$

Q12 Not graded

2. In trying to decide whether or not to go hiking, some of the factors I consider are the weather and whether or not I have chores to do at home. There are other factors as well of course. Some example decisions are summarized in the table below.

Weather	ChoresToDo	Hike
Sunny	chores	yes
Sunny	chores	yes
Sunny	no chores	yes
Sunny	no chores	yes
Sunny	no chores	yes
Sunny	no chores	yes
Cloudy	chores	no
Cloudy	chores	no
Cloudy	no chores	yes
Cloudy	no chores	no
Cloudy	chores	yes
Cloudy	chores	yes

With 9 positive examples of hiking and 3 negatives, the entropy or information of this decision in bits is 0.81. Consider the following decision trees, reflecting splitting on the attributes of Weather or ChoresToDo:



Q13 Score: **11**

8E58D5EE-5D66-4B42-ADA1-0F3D98C4EB93

SPCSCI561-examFinalg

#27 14 of 19

a) The ChoresToDo tree has been filled out. Complete the values for the Weather tree, including entropy. (Show your work) (6 points)

sunny: Yes 6 no: 0 Entropy = $I(\frac{6}{6}, 0) = 0$

cloudy: Yes 3, no: 3 Entropy = $I(\frac{1}{2}, \frac{1}{2}) = 1$

b) Calculate the information gain *IG* from splitting on ChoresToDo and Weather. Please show formulas used and steps clearly. You do not need to compute a final numerical result. Plug in values to the formulas so that the result could be calculated with a calculator. (4 points)

IG(ChoresToDo) = $I(\frac{9}{12}, \frac{3}{12}) - \left(\frac{6}{12} I(\frac{4}{6}, \frac{2}{6}) + \frac{6}{12} I(\frac{5}{6}, \frac{1}{6})\right)$

IG(Weather) = $I(\frac{9}{12}, \frac{3}{12}) - \left(\frac{6}{12} I(\frac{6}{6}, 0) + \frac{6}{12} I(\frac{1}{2}, \frac{1}{2})\right)$

c) Which of the two attributes is a better choice in constructing a decision tree? Why? (2 points)

weather is a better choice because it is more simple

reason is incorrect Correct reason : Because the information gain (IG) from it is more.

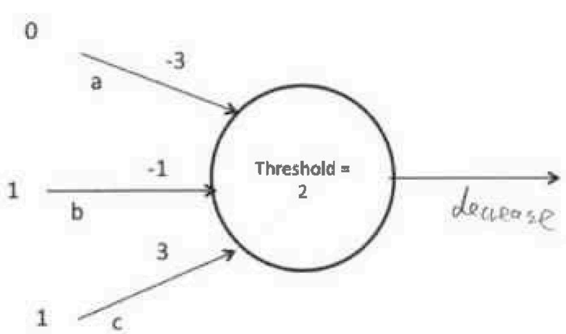
Q14 Score: **2**

7425F515-D878-449A-8E0B-F0CA0739ADEF

SPCSCI561-examFinalg

#27 15 of 19

6. Perceptrons and Neural Nets (16 points)



x	0	1
y	0	0
z	0	0
Out	0	

1. Consider the perceptron shown above. Assume that its firing follows the threshold rule we saw in class.

a) With the inputs and weights as shown, what would be the output of the perceptron? (1 point) 1

b) If the output from the above case was higher than the desired value, how would each of the weights on inputs a, b, and c be adjusted? Answers should be "increase", "decrease" or "stay the same" (3 points)

a? increase

b? stay the same.

c? decrease

Q15 Score: **4**

220290F7-123A-426E-881E-96665037F7D1

SPCSCI561-examFinalg

#27 16 of 19

2. You are building a set of perceptrons to mimic logic circuits, using logical operators such as AND, OR, NOT. Assume that your perceptrons have a threshold function that produces 1 if the threshold is equaled or exceeded and 0 otherwise. Assume all inputs have value either 0 or 1.

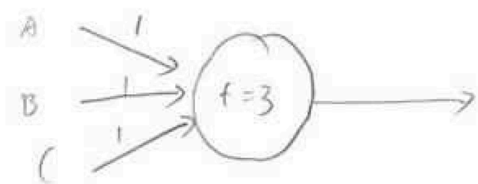
a) Show how perceptrons could be used to mimic the following logical expressions. For each case draw the perceptron(s), show the weights on inputs and the thresholds.

1) Use a **single** layer perceptron to implement: (4 points)

$A \wedge B \wedge C$

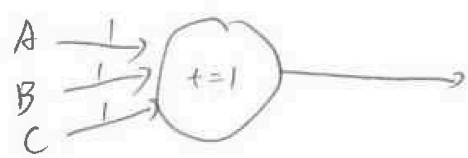
and

A B C → 1
1 1 1
→ 0



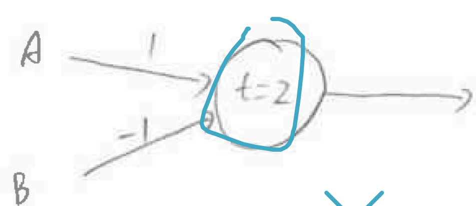
$A \vee B \vee C$

A B C → 0
0 0 0
→ 1



2) Use one or more layers to implement $A \wedge \neg B$ (3 points)

A∧¬ B
1 0 → 1
→ 0.



Q16 Score: **3**

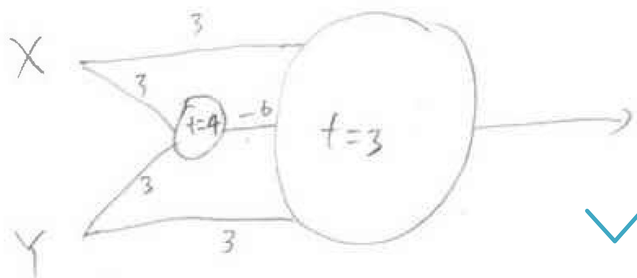
b) In class we mentioned that a single layer perceptron cannot implement XOR but a multilayer perceptron can.

1) Why not? (2 points)

XOR 00, 11 → 0 01, 10 → 1

Only one layer cannot make 11 to 0.

2) Show a multilayer perceptron implementation of XOR. (3 points)



Q17 Score: 0

3A283356-AA28-4BC4-9E31-2C3B7667CDAA

SPCSCI561-examFinalg

#27 18 of 19

7. Learning – Naïve Bayes (17 points)

1. Given the following training data:

TEXT	CLASS
Hello hope your exam went fine	HAM
Credit card details required	SPAM
Signup to get free credit card	SPAM
Win a free trip to Paris	SPAM
Hello exam went fine	HAM

a) Using add one smoothing what is: (4 points)

1) P(win | SPAM)?

$= \frac{2}{17}$ $\frac{P(\text{win spam})}{P(\text{spam})}$ $\frac{3}{3}$

2) P(my | HAM)?

$= \frac{1}{17}$

b) Given the message: " Win Your Free Card ". How do you calculate P(SPAM | msg)? Give formulas for every step involved. *Do not use add one smoothing for this part.* (5 points)

Q18 Score: **3**

EC4A977C-61C4-47CF-88BA-87AECA4993C4

SPCSCI561-examFinalg

#27 19 of 19

2. In the Bag of Words representation, are the following represented the same? (1 point)
   A. Under the table, there were lots of toys!
   B. Were there lots of toys under the table?

Yes, do not care order.

3. The actual and predicted outcomes of a classifier are given below:

Actual	Predicted
yes	yes
yes	yes
no	yes
yes	yes
yes	yes
no	yes
yes	yes
yes	yes
yes	yes
yes	yes
no	no

a) What is the precision of the classifier for class "**yes**"? (show your work) (2 points)

0.8%

work not shown

1

b) What is the recall for class "**no**"? (show your work) (2 points)

$P = \frac{1}{3}$

work not shown

1

c) What is the "F-score" for class "yes"? (show your work) (3 points)

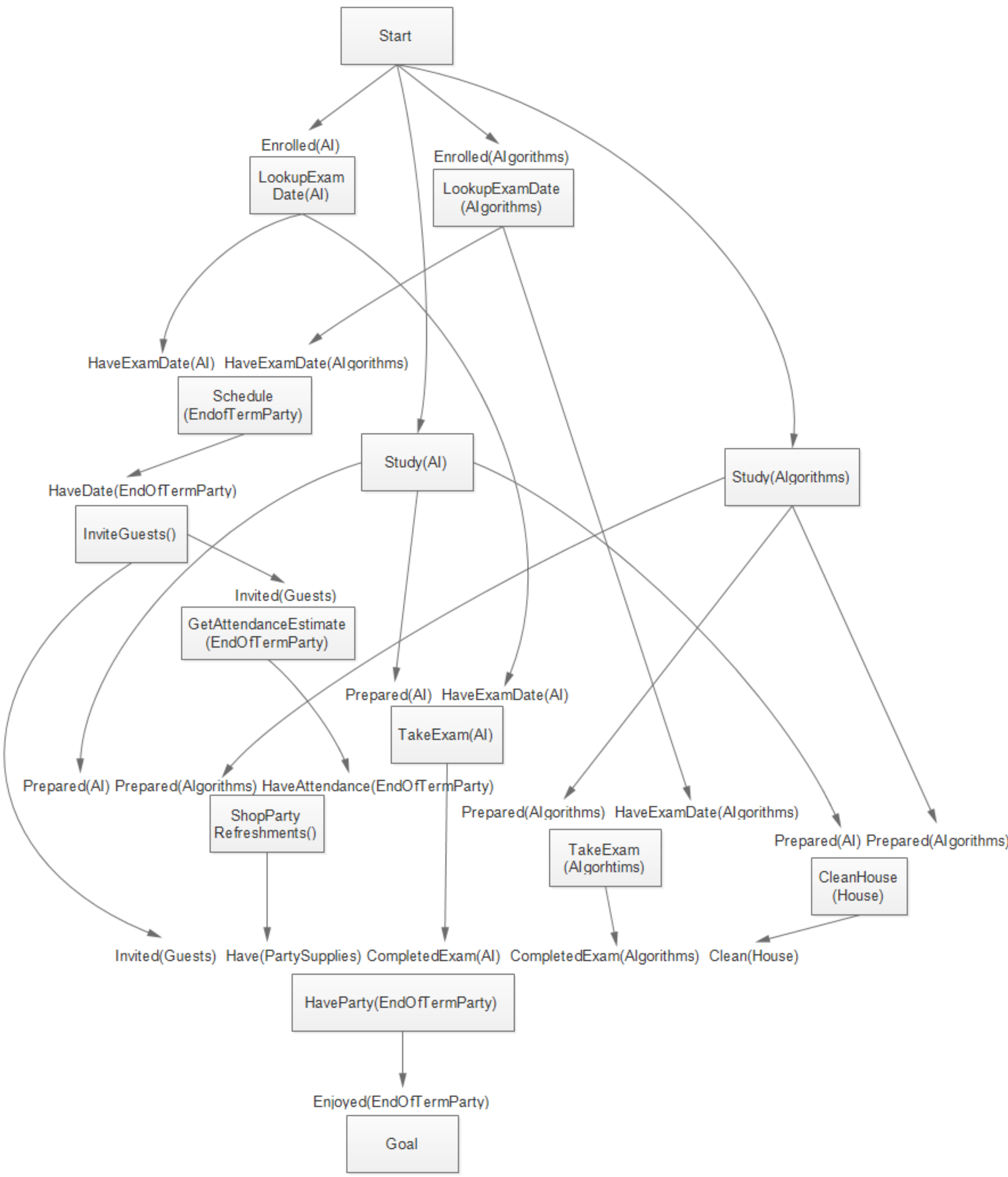
1- Multiple Choice Questions

1) d   2) a   3) f   4) d   5) e

2- True/False Questions

1) F   2) F   3) T   4) T   6) F   7) T   8) F   9) F   10) T

3- Planning



4- Bayesian Probabilities

1) a) P (strep| fever) = $\frac{P\ (fever \mid strep\ ) * P\ (strep\ )}{P(fever)}$ = $\frac{P\ (fever \mid strep\ ) * P\ (strep\ )}{P(fever\ \ |strep\ )*P(strep\ )+P(fever \mid \sim strep\ )*P(\sim strep)}$ =

$\frac{0.6*0.15}{0.6*0.15+0.3*0.85}$ = 6/23 = 0.261

b) P(strep|~fever) = $\frac{P\ (\sim fever \mid strep\ ) * P\ (strep\ )}{P(\sim fever)}$ =

$\frac{P\ (\sim fever \mid strep\ ) * P\ (strep\ )}{P(\sim fever\ \ |strep\ )*P(strep\ )+P(\sim fever \mid \sim strep\ )*P(\sim strep)}$ = $\frac{0.4*0.15}{0.4*0.15+0.7*0.85}$ = 12/131 =0.092

2) a) P(strep | fever, test) = $\frac{P\ (fever,test \mid strep\ ) * P\ (strep\ )}{P(fever,test)}$ =

$\frac{P\ (fever \mid strep\ )*P(test|\ strep)*\ P\ (strep\ )}{P(fever|strep)P(test\ |strep)P(strep)+P(fever|\sim strep)P(\sim strep)P(\sim strep)}$ = $\frac{0.6*0.95*0.15}{0.6*0.95*0.15+0.3*0.1*0.85}$ =57/74 =0.770

b) P(strep | fever, ~test) = $\frac{P\ (fever,\sim test \mid strep\ ) * P\ (strep\ )}{P(fever,\sim test)}$ =

$\frac{P\ (fever \mid strep\ )*P(\sim test|\ strep)*\ P\ (strep\ )}{P(fever|strep)P(\sim test\ |strep)P(strep)+P(fever|\sim strep)P\ (\sim strep)P(\sim strep)}$ = $\frac{0.6*0.05*0.15}{0.6*0.05*0.15+0.3*0.9*0.85}$ =1/52=0.019

3) a) P(S,E,K,~A) = P(S) * P(E|S) * P(K|S) * P(~A|E,K) = 0.7 * 0.3 * 0.8 * 0.1 = 0.0168

b) P(K | S, E) = P (K | S) = 0.8

c) P(S | A) = $\frac{P(S,A)}{P(A)}$ = = $\frac{\sum_K \sum_E P(S,A,K,E)}{\sum_K \sum_E \sum_S P(S,A,K,E)}$ =

$\frac{P(S,A,K,E)+P(S,A,K,\sim E)+P(S,A,\sim K,E)+P(S,A,\sim K,\sim E)}{P(S,A,K,E)+P(S,A,K,\sim E)+P(S,A,\sim K,E)+P(S,A,\sim K,\sim E)+\ P(\sim S,A,K,E)+P(\sim S,A,K,\sim E)+P(\sim S,A,\sim K,E)+P(\sim S,A,\sim K,\sim E)}$

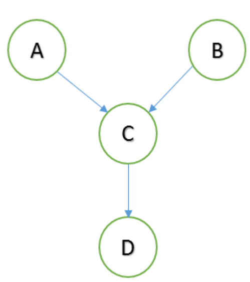
Nominator = 0.7*0.9*0.8*0.3+0.7*0.5*0.8*0.7+0.7*0.6*0.2*0.3+0.7*0.1*0.2*0.7

Denominator = Nominator

+0.3*0.9*0.4*0.8+0.3*0.5*0.4*0.2+0.3*0.4*0.6*0.8+0.3*0.1*0.6*0.2

4) a) A : 1 , B : 1 , C : 2 , D : 4

b)



5) **Decision Tree Learning**

1) I(P) = I ($\frac{12}{16}$, $\frac{4}{16}$) = $\frac{-3}{4}$ log($\frac{3}{4}$) - $\frac{1}{4}$ log($\frac{1}{4}$) = 0.81 (not wanted)

2) a) weather: sunny → 6 yes , 0 no   entropy = -1 * log(1) - 0*log(0) = 0

weather: cloudy → 3 yes , 3 no   entropy = -0.5*log(0.5) – 0.5*log(0.5) = 1

b) IG (choresToDo) = I (Hike) – remainder (choresToDo) =

$0.81 - [\frac{1}{2} * I(\frac{4}{6}, \frac{2}{6}) + \frac{1}{2} * I(\frac{5}{6}, \frac{1}{6})] = 0.81 - [0.5*0.92 + 0.5*0.65] = 0.025$

IG(weather) = I (Hike) – remainder(weather) =

$0.81 - [\frac{1}{2} * I(1,0) + \frac{1}{2} * I(0.5,0.5)] = 0.81 - 0.5 = 0.31$

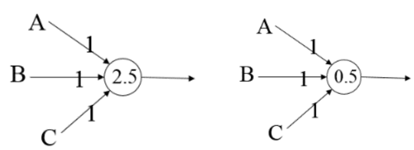
c) weather is the better choice. Because the information gain (IG) from it is more.

6) Perceptrons and Neural Nets

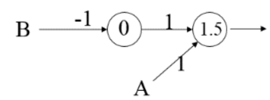
1) a) 1

b) a: stay the same , b: decrease c: decrease

2) a) 1)

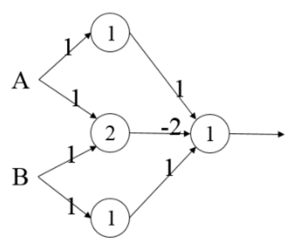


2)



b) 1) a single layer perceptron can only represent linearly separable functions. XOR is not a linearly separable function.

2)



7) Learning – Naïve Bayes

1) a) 1) P(win | SPAM) = $\frac{n+1}{N+k} = \frac{1+1}{16+18} = \frac{1}{17}$ 2) P (my | HAM) = $\frac{n+1}{N+k} = \frac{0+1}{+18} = 1$

b) P(SPAM | msg) = $\frac{P(msg|SPAM)\,P(SPAM)}{P(msg|SPAM)\,P(SPAM) + P(msg|HAM)\,P(HAM)}$

P(msg | SPAM) = P(win|SPAM) * P(your|SPAM) * P(free|SPAM) * P(card|SPAM)

P(msg | HAM) = P(win|HAM) * P(your|HAM) * P(free|HAM) * P(card|HAM)

P(SPAM | msg) = $\frac{\frac{1}{16} * \frac{0}{16} * \frac{2}{16} * \frac{2}{16} * \frac{3}{5}}{\frac{1}{16} * \frac{0}{16} * \frac{2}{16} * \frac{2}{16} * \frac{3}{5} + \frac{0}{10} * \frac{1}{10} * \frac{0}{10} * \frac{0}{10} * \frac{2}{5}} = 0$

2) yes

3) a) Precision(yes) = $\frac{count(\textit{correctly classified as yes})}{count(\textit{classified as yes})} = \frac{8}{10}$

b) Recall(no) = $\frac{count(correctly\ classified\ as\ no)}{count(belongs\ in\ no)}$ = $\frac{1}{3}$

c) F = $\frac{2*precision*recall}{precision+recall}$ = $\frac{2*\frac{8}{10}*\frac{8}{8}}{\frac{8}{10}+\frac{8}{8}}$ = $\frac{8}{9}$