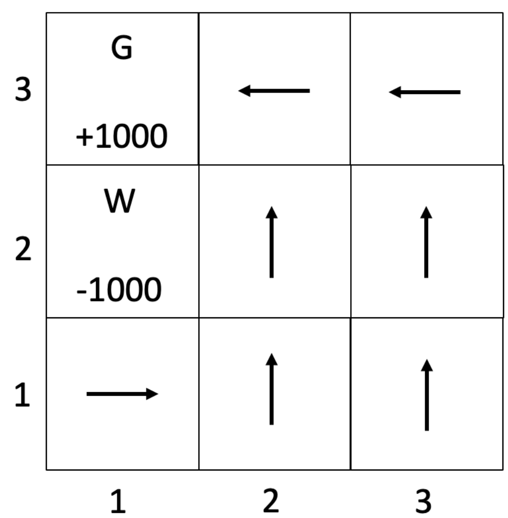
Washington State University
School of Electrical Engineering and Computer Science
Fall 2018

CptS 440/540 Artificial Intelligence
Homework 12 Solution
Due: December 6, 2018 (11:59pm)

General Instructions: Put your answers to the following problems into a PDF document and submit as an attachment under Content → Homework 12 for the course CptS 440 Pullman (all sections of CptS 440 and 540 are merged under the CptS 440 Pullman section) on the Blackboard Learn system by the above deadline. Note that you may submit multiple times, but we will only grade the most recent entry submitted before the above deadline.

1. Consider the 3x3 wumpus world shown below. The goal of this simplified game is to be collocated with the gold (where we get a +1000 reward) and not collocated with the wumpus (or we get a -1000 reward). All other states have a reward of -1. As before, the agent starts in [1,1], but has only four possible actions: Up, Down, Left, Right (there is no orientation or turning). Each of these actions always works, although attempting to move into a wall results in the agent not moving. We will use reinforcement learning to solve this problem.



   a. Compute the utility *U(s)* of each non-terminal state *s* given the policy shown above. Note that [1,2] and [1,3] are terminal states, where U([1,2]) = -1000, and U([1,3]) = +1000. You may assume $\gamma = 0.9$.
   b. Using temporal difference Q-learning, compute the Q values for Q([1,1],Right), Q([2,1],Up), Q([2,2],Up), Q([2,3],Left), after each of five executions of the action sequence: Right, Up, Up, Left (starting from [1,1] for each sequence). You may assume $\alpha = 1$, $\gamma = 0.9$, and all Q values for non-terminal states are initially zero.

Solution:

a. Below are the equations for the utility values. These can be solved analytically or iteratively.

$U(1,2) = -1000$

$U(1,3) = +1000$

$U(2,3) = R(2,3) + \gamma\, U(1,3) = (-1) + (0.9)*U(1,3) = (-1) + (0.9)*(1000) = 899.0$

$U(2,2) = R(2,2) + \gamma\, U(2,3) = (-1) + (0.9)*U(2,3) = (-1) + (0.9)*(899) = 808.1$

$U(2,1) = R(2,1) + \gamma\, U(2,2) = (-1) + (0.9)*U(2,2) = (-1) + (0.9)*(808.1) = 726.3$

$U(1,1) = R(1,1) + \gamma\, U(2,1) = (-1) + (0.9)*U(2,1) = (-1) + (0.9)*(726.3) = 652.7$

$U(3,3) = R(3,3) + \gamma\, U(2,3) = (-1) + (0.9)*U(2,3) = (-1) + (0.9)*(899) = 808.1$

$U(3,2) = R(3,2) + \gamma\, U(3,3) = (-1) + (0.9)*U(3,3) = (-1) + (0.9)*(808.1) = 726.3$

$U(3,1) = R(3,1) + \gamma\, U(3,2) = (-1) + (0.9)*U(3,2) = (-1) + (0.9)*(726.3) = 652.7$

b. Below are the equations for the Q values followed by a table showing their values after each iteration. Note that $Q(1,3,a)=1000$ for all actions a.

$Q(1,1,right) = Q(1,1,right) + \alpha\, (R(1,1) + \gamma \max_a Q(2,1,a) - Q(1,1,right))$

$Q(2,1,up) = Q(2,1,up) + \alpha\, (R(2,1) + \gamma \max_a Q(2,2,a) - Q(2,1,up))$

$Q(2,2,up) = Q(2,2,up) + \alpha\, (R(2,2) + \gamma \max_a Q(2,3,a) - Q(2,2,up))$

$Q(2,3,left) = Q(2,3,left) + \alpha\, (R(2,3) + \gamma \max_a Q(1,3,a) - Q(2,3,left))$

Iteration	Q(1,1,right)	Q(2,1,up)	Q(2,2,up)	Q(2,3,left)
0	0	0	0	0
1	-1	-1	-1	899
2	-1	-1	808.1	899
3	-1	726.3	808.1	899
4	652.7	726.3	808.1	899
5	652.7	726.3	808.1	899

2. Suppose you are given a set of 1,000 messages sent from the Wumpus, where 600 of the messages are classified as describing the Wumpus as hungry, and 400 of the messages are classified as describing the Wumpus as not hungry. After analysis of the messages, you compute the following bigram models for the two classes: hungry and not hungry.

word1	**word2**	**hungry**	**not hungry**
the	wumpus	300	150
wumpus	is	90	40
is	not	60	60
not	hungry	100	200
is	hungry	300	200
is	full	100	200
is	starving	50	100
the	agent	150	200
agent	is	40	90
not	full	100	50
not	starving	50	100

You receive message *m* = "the wumpus is not hungry" and want to determine which class this message belongs to: hungry or not hungry. Do the following:

a. Compute the probabilities P(Class = hungry) and P(Class=not hungry). Show your work.
b. Compute the probability P(Class = hungry | Message = *m*) using the bigram model. Show your work.
c. Compute the probability P(Class = not hungry | Message = *m*) using the bigram model. Show your work.
d. Which class is more likely for message *m*?

Solution

a. P(Class = hungry) = 600 / 1000 = 0.6
   P(Class = not hungry) = 400 / 1000 = 0.4

b. P(Class = hungry | Message = "the wumpus is not hungry")
   = α P(Message = "the wumpus is not hungry" | Class = hungry) * P(Class = hungry)
   = α P("wumpus" | "the")*P("is" | "wumpus")*P("not" | "is")*P("hungry" | "not")*P(Class = hungry)
   Note that P(word2 | word1) = P(word1 word2) / P(word1) = [Freq(word1 word2) / Freq(all words)] / [Freq(word1 <anyword>) / Freq(all words)] = Freq(word1 word2) / Freq(word1 ?)
   Using frequencies from the bigram model for Class = hungry:
   = α (300/(300+150))*(90/90)*(60/(60+300+100+50))*(100/(100+100+50))*(0.6)
   = α (0.0188)
   Based on this result and result from (c), α = 34.11
   = 0.64

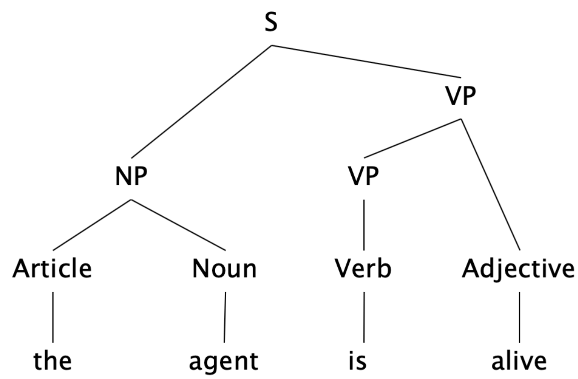
c. P(Class = not hungry | Message = "the wumpus is not hungry")
   = α P(Message = "the wumpus is not hungry" | Class = not hungry) * P(Class = not hungry)
   = α P("wumpus" | "the")*P("is" | "wumpus")*P("not" | "is")*P("hungry"|"not")*P(Class=not hungry)
   Using frequencies from the bigram model for Class = not hungry:
   = α (150/(150+200))*(40/40)*(60/(60+200+200+100))*(200/(200+50+100))*(0.4)
   = α (0.0105)
   Based on this result and result from (b), α = 1/(0.0188+0.0105) = 34.11
   = 0.36

d. Class=hungry is more likely (can't trust a hungry Wumpus ☺).

3. Based on the Wumpus world lexicon and grammar on the last page, show the parse trees for the following sentences. Note that it is possible for a sentence to have no parse or more than one parse tree (in which case show them all).
   a. "the agent is alive"
   b. "the agent is near the wumpus and the gold"
   c. "the agent shoots the wumpus in 1 3"
   d. "the wumpus who stinks is dead"

Solution:

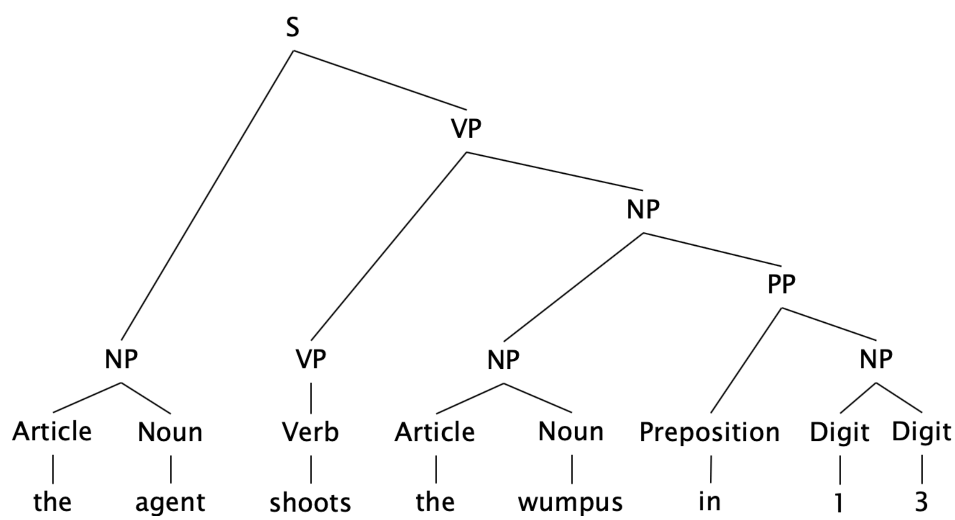
a. "the agent is alive"

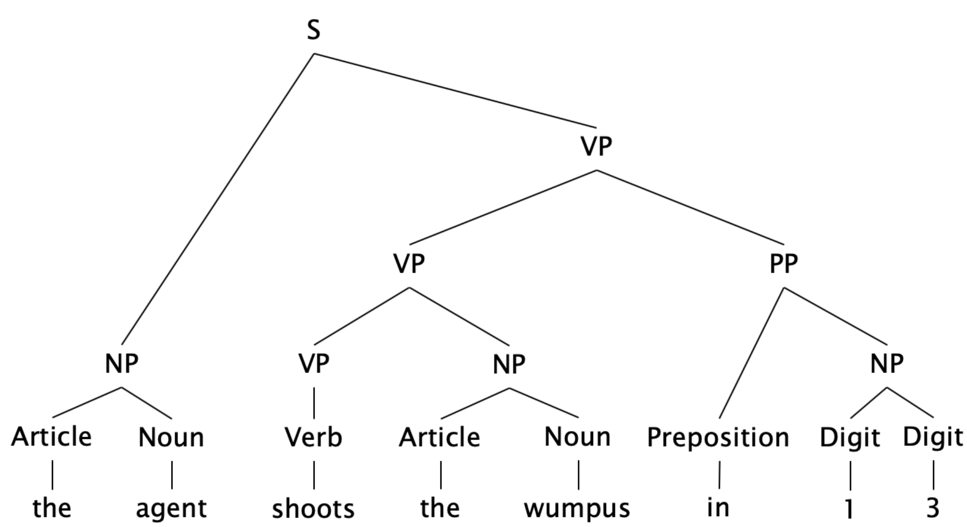


b. "the agent is near the wumpus and the gold"

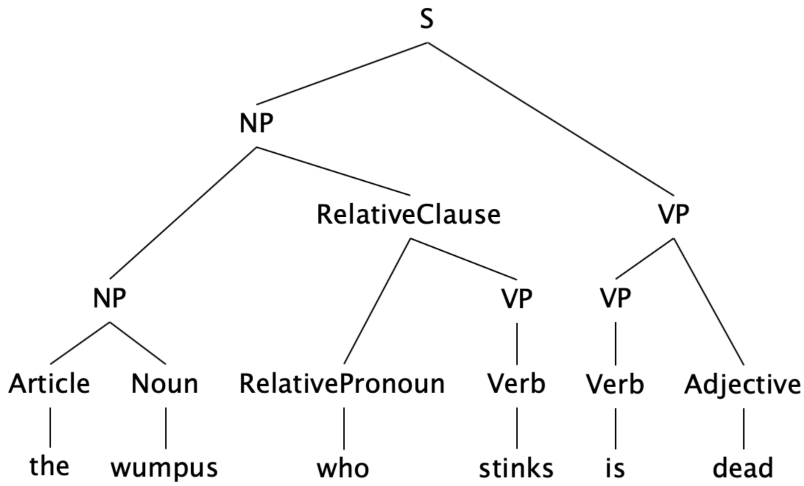
No parse. The grammar does accept conjunctive sentences, but not conjunctive noun phrases.

c. "the agent shoots the wumpus in 1 3" (2 parses)





d. “the wumpus who stinks is dead”



4. The Hidden Markov Model (HMM) for the [hh] phoneme is shown on the last page. Calculate the most probable path through this HMM for the output sequence [C_2,C_3,C_3,C_4,C_5,C_6,C_7]. Also give its probability. Show your work.

   Solution:

   There are six possible paths through the HMM consistent with the output sequence. Each of these paths are shown below with the probability.

   Onset, Onset, Onset, Mid, Mid, End, End, Final
   (0.5) (0.6)(0.3) (0.6)(0.3) (0.4)(0.3) (0.5)(0.4) (0.5)(0.4) (0.3)(0.4) (0.7) = 6.53×10^{-6}

   Onset, Onset, Onset, Mid, End, End, End, Final
   (0.5) (0.6)(0.3) (0.6)(0.3) (0.4)(0.3) (0.5)(0.2) (0.3)(0.4) (0.3)(0.4) (0.7) = 1.96×10^{-6}

   Onset, Onset, Mid, Mid, Mid, End, End, Final
   (0.5) (0.6)(0.3) (0.4)(0.3) (0.5)(0.3) (0.5)(0.4) (0.5)(0.4) (0.3)(0.4) (0.7) = 5.44×10^{-6}

   Onset, Onset, Mid, Mid, End, End, End, Final
   (0.5) (0.6)(0.3) (0.4)(0.3) (0.5)(0.3) (0.5)(0.2) (0.3)(0.4) (0.3)(0.4) (0.7) = 1.63×10^{-6}

   Onset, Mid, Mid, Mid, Mid, End, End, Final
   (0.5) (0.4)(0.3) (0.5)(0.3) (0.5)(0.3) (0.5)(0.4) (0.5)(0.4) (0.3)(0.4) (0.7) = 4.54×10^{-6}

   Onset, Mid, Mid, Mid, End, End, End, Final
   (0.5) (0.4)(0.3) (0.5)(0.3) (0.5)(0.3) (0.5)(0.2) (0.3)(0.4) (0.3)(0.4) (0.7) = 1.36×10^{-6}

   The most likely path is Onset, Onset, Onset, Mid, Mid, End, End, Final with probability 6.53×10^{-6}.

Or, we can use the Viterbi algorithm:

$V_{1,Onset} = P(C_2|Onset)P(Onset) = (0.5)(1.0) = 0.5$
$V_{1,Mid} = V_{1,End} = V_{1,Final} = 0$

$V_{2,Onset} = P(C_3|Onset)\max\{[P(Onset|Onset)V_{1,\underline{Onset}}], 0, 0, 0\} = (0.3)\{(0.6)(0.5)\} = 0.09$
$V_{2,Mid} = P(C_3|Mid)\max\{[P(Mid|Onset)V_{1,Onset}], 0, 0, 0\} = (0.3)\{(0.4)(0.5)\} = 0.06$
$V_{2,End} = V_{2,Final} = 0$

$V_{3,Onset} = P(C_3|Onset)\max\{[P(Onset|Onset)V_{2,\underline{Onset}}], 0, 0, 0\} = (0.3)\{(0.6)(0.09)\} = 0.0162$
$V_{3,Mid} = P(C_3|Mid)\max\{[P(Mid|Onset)V_{2,Onset}], [P(Mid|Mid)V_{2,Mid}], 0, 0\}$
$= (0.3)\max\{(0.4)(0.09), (0.5)(0.06), 0, 0\} = (0.3)\max\{0.036, 0.03, 0, 0\} = 0.0108$
$V_{3,End} = V_{3,Final} = 0$

$V_{4,Onset} = 0$
$V_{4,Mid} = P(C_4|Mid)\max\{[P(Mid|Onset)V_{3,\underline{Onset}}], [P(Mid|Mid)V_{3,Mid}], 0, 0\}$
$= (0.3)\max\{(0.4)(0.0162), (0.5)(0.0108), 0, 0\}$
$= (0.3)\max\{0.00648, 0.0054, 0, 0\} = 0.001944$
$V_{4,End} = V_{4,Final} = 0$

$V_{5,Onset} = 0$
$V_{5,Mid} = P(C_5|Mid)\max\{0, [P(Mid|Mid)V_{4,\underline{Mid}}], 0, 0\} = (0.4)\{(0.5)(0.001944)\} = 0.0003888$
$V_{5,End} = P(C_5|End)\max\{0, [P(End|Mid)V_{4,Mid}], 0, 0\} = (0.2)\{(0.5)(0.001944)\} = 0.0001944$
$V_{5,Final} = 0$

$V_{6,Onset} = V_{6,Mid} = 0$
$V_{6,End} = P(C_6|End)\max\{0, [P(End|Mid)V_{5,\underline{Mid}}], [P(End|End)V_{5,End}], 0\}$
$= (0.4)\max\{0, (0.5)(0.0003888), (0.3)(0.0001944), 0\}$
$= (0.4)\max\{0, 0.0001944, 0.00005832, 0\} = 0.00007776$
$V_{6,Final} = 0$

$V_{7,Onset} = V_{7,Mid} = 0$
$V_{7,End} = P(C_7|End)\max\{0, 0, [P(End|End)V_{6,\underline{End}}], 0\} = (0.4)\{(0.3)(0.00007776)\} = 9.3312 \times 10^{-6}$
$V_{7,Final} = 0$

The final probability is $P(Final|End)V_{7,\underline{End}} = (0.7)(9.3312 \times 10^{-6}) = 6.53 \times 10^{-6}$. The most probable path is comprised of the states of the V's used in the maximum at each stage (underlined in red above): Onset, Onset, Onset, Mid, Mid, End, End, Final.

Lexicon:

S → NP VP | S Conjunction S

NP → Pronoun | Noun | Article Noun | Article Adjectives Noun
| Digit Digit | NP PP | NP RelativeClause

VP → Verb | VP NP | VP Adjective | VP PP | VP Adverb

Adjectives → Adjective | Adjective Adjectives

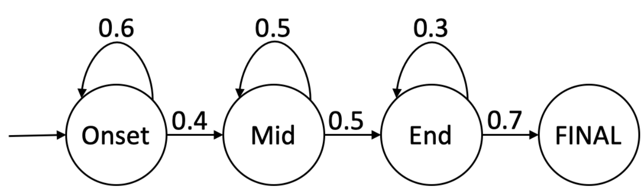
PP → Preposition NP

RelativeClause → RelativePronoun VP

Grammar:

Noun → stench | breeze | glitter | wumpus | pit | agent | gold | arrow
Verb → is | sees | smells | shoots | feels | stinks | grabs | eats
Adjective → right | left | smelly | breezy | alive | dead
Adverb → here | there | near | ahead
Pronoun → me | you | I | it
RelativePronoun → that | which | who | whom
Article → the | a | an | every
Preposition → to | in | on | of | near
Conjunction → and | or | but | yet
Digit → 0| 1| 2| 3| 4| 5| 6| 7| 8| 9

Phone HMM for [hh]:



Output probabilities for the phone HMM:

Onset:	Mid:	End:
C_1: 0.2	C_3: 0.3	C_5: 0.2
C_2: 0.5	C_4: 0.3	C_6: 0.4
C_3: 0.3	C_5: 0.4	C_7: 0.4