Shrina Parikh

Prof. Davis

AI Final Exam

MC

1. C - KNN
2. A - vector of weights on the arcs in the networks
3. F - $P(x,y|z) = P(x|z) \cdot P(y|z)$
4. A - $\forall_p F(p,S) \Rightarrow \exists_m P(p,m) \wedge \neg L(S,m)$
5. D - The number of words in the language

~~i for doc. D matrix vector~~

Part II

6.
A. weights
B. hyperparameter
C. training set
D. labelled instance
E. in the set $\{-1, 0, 1\}$
F. nature of the classification error

→ ~~8 to E~~

7.
A. predictive attributes
B. training set
C. feed forward
D. output layer
E. Classification attribute
F. back propagation
G. output layer to the input layer
H. weights

8.
A. A and B are tags
B. a word and a number
C. word Z
D. tag C

Part III

9. A. flat - non-hierarchal structure of clusters
hard - a data point can only belong to 1 cluster
disjoint - there are no overlapping elements between clusters
exhaustive - every point belongs to a cluster

soft clustering algo: EM (Expectation Maximization)
non-flat clustering algo: agglomerative clustering
↳ reaches binary hierarchy of clusters

B.

Cluster 1: A, B, C	avg: 1.5	$(1.5)^2 = 2.25$	$2(0.5)^2 = 0.5$
cluster 2: D, E, F	avg: 24	$2^2 + 2^2$	
cluster 3: G, H	avg: 33	$1^2 + 1^2$	

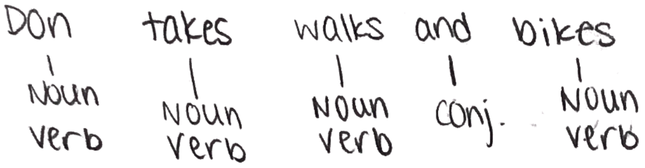
C. The resulting value from the objective function is the sum of the differences squared b/w each point in the cluster and the average of the points in that cluster.
Numerical val for pt. B: $\underbrace{2.25 + 0.5 + 8 + 2 = 12.75}_{\text{work above}}$

D. {A, D} {B, C} {E, F, G, H}
avg/center: 11 1.5 29
running 1-iter k-means
↓ ↓ ↓
{ } {A, B, C} {D, E, F, G, H}
new avg: 1.5 new avg: 27.6
Outcome
Stops after this to do random restart

Of the values 11, 1.5, and 29, A is closest to 1.5 so it moves + D is closest to 29 so it moves also

E. It illustrates how k-means is vulnerable to starvation as we get an empty cluster.
we end with less than k clusters.
Soln: random restart

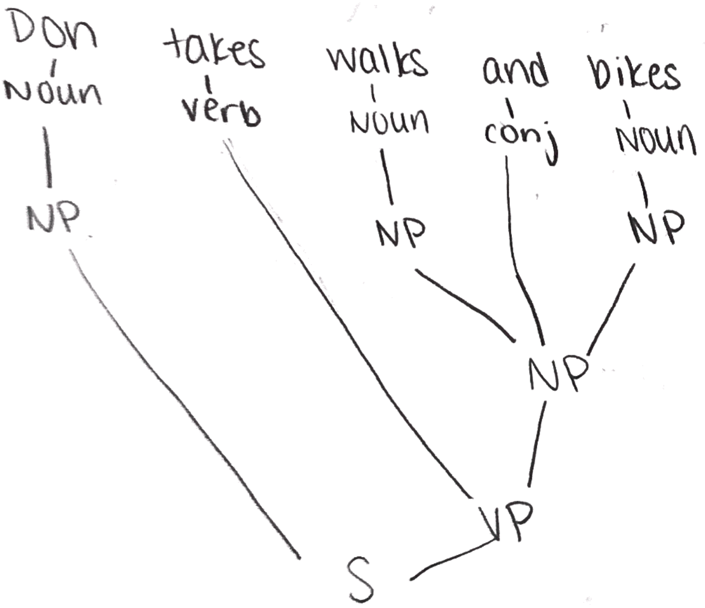
10

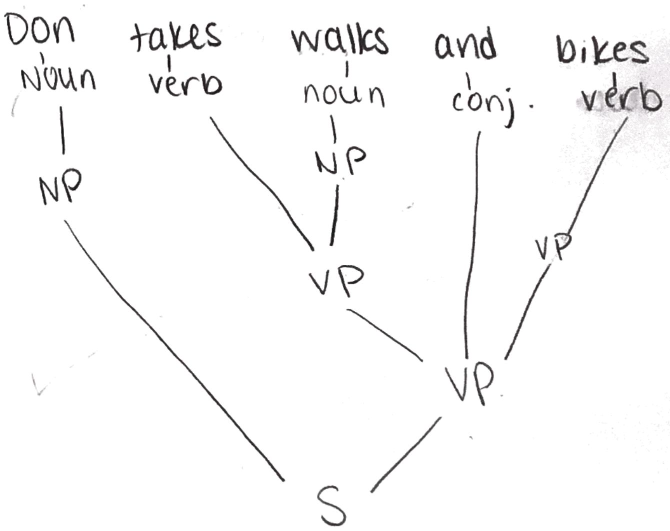


S = NP VP ✓

~~S = S conj. S~~

not possible





~~Don takes walks and bikes~~

~~conj.~~

11. A. Blue $C = 0.7(B) + 0.3(0)$

$= 0.7(-0.371)$

$= -0.259$

$-0.259 > -0.46$ so the blue action will be chosen in updating the policy for C

B. $v(A) = 0.8(C) + 0.2(-1)$

$v(B) = 0.1(1) + 0.9(A)$

$v(C) = 0.7(B) + 0.3(0)$

(12) a) Sometimes 0 values can enter the product if $\#_T(X.C=v, X.A_j=u)=0$, so $Freq_T(X.A_j=u \mid X.C=v)=0$, making the product 0

This single attribute then vetoes the value $X.C=v$ even if other attributes favor $X.C=v$

Laplacian correction is used so 1 attribute does not cause others to be disregarded in case of 0 values

b) Formula for correction:

$$P(X.A_i=u \mid X.C=v) = \frac{\#(X.A_i=u, X.C=v)+\boxed{\delta} \rightarrow \text{never } 0}{\#(X.C=v)+\boxed{q\delta} \rightarrow \text{greater weight b/c } q \text{ as coeff. if mult. values of } A_j}$$

where δ is a small number (metaparameter)

q = # of different values of Aj

A_j = attribute w/ 0 freq given $X.C=v$

C = classification attribute

This solves the problem because the resulting probability will never be 0.

13 HA → HB adj.

has peg empty ⇒ Put peg in HB

A. $Put(9,2) \Rightarrow Peg(8,2) \wedge -Peg(9,2)$

HA = 8
HB = 9

B. $Put(9,2) \Rightarrow Peg(8,3) \wedge Peg(9,3)$

C. $\neg Peg(6,4) \wedge Peg(6,5) \Rightarrow Jump(1,3,6,4) \vee Jump(4,5,6,4) \vee \underline{\underline{Put(6,4)}}$