

## 2. Given these conditional probabilities

 $P(Mary)=0.1; \quad P(likes)=0.2; \quad P(cats)=0.3 \, . \quad P(Mary|likes)=0.2; \quad P(likes|Mary)=0.3; \quad P(cats|likes)=0.1; \\ P(likes|cats)=0.4$ 

Approximate the probability of the following sentence with bigrams: "Mary likes cats"

- P(Mary likes cats) = 0.003
- O P(Mary likes cats) = 0.008
- O P(Mary likes cats) = 0
- O P(Mary likes cats) =1

Correct
Correct.

-	C: +l		and the least that are
ა.	Given these	conditional	probabilities

P(Mary)=0.1; P(likes)=0.2; P(cats)=0.3

P(Mary|<s>)=0.2; P(</s>|cats)=0.6

P(likes|Mary) =0.3; P(cats|likes)=0.1

Approximate the probability of the following sentence with bigrams: "<s> Mary likes cats </s>"

- $\bigcirc$  P(<s> Mary likes cats </s>) = 0.003
- $\bigcirc$  P(<s> Mary likes cats </s>) = 1
- $\bigcirc$  P(<s> Mary likes cats </s>) = 0
- P(<s> Mary likes cats </s>) = 0.0036

**⊘** Correct

1	1	poin
- /	-	P

**4.** Given the logarithm of these conditional probabilities:

log(P(Mary|<s>))=-2; log(P(</s>|cats))=-1

log(P(likes|Mary)) =-10; log(P(cats|likes))=-100

Approximate the log probability of the following sentence with bigrams: "<s> Mary likes cats </s>"

- log(P(<s> Mary likes cats </s>)) = -113
- $\bigcirc$  log(P(<s> Mary likes cats </s>)) = 2000
- $\bigcirc$  log(P(<s> Mary likes cats </s>)) = -112
- O log(P(<s> Mary likes cats </s>)) = 113

Correct
Correct

<ol><li>Given the logarithm of these conditional probabilit</li></ol>	ies:
---	------

 $log(P(Mary|<\!\!s\!\!>))=\!\!-2;\quad log(P(<\!\!/s\!\!>|cats))=\!\!-1$ 

log(P(likes|Mary)) =-10; log(P(cats|likes))=-100

Assuming our test set is W="<s> Mary likes cats </s>", what is the model's perplexity.

- $\bigcirc$  log PP(W) = -113
- O  $\log PP(W) = (-1/5)*(-113)$
- log PP(W) = (-1/4)\*(-113)
- O  $\log PP(W) = (-1/5)*113$

**⊘** Correct

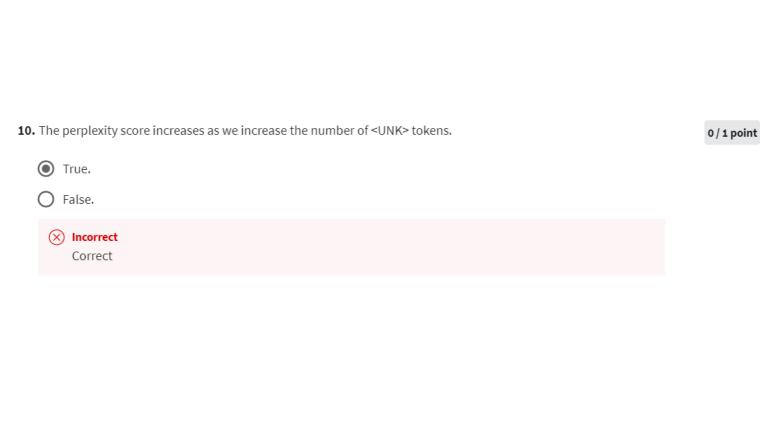
Correct.

6.	Given the training corpus and minimum word frequency=2, how would the vocabulary for corpus preprocessed with <unk> look like?</unk>	1 / 1 pc
	" <s>I am happy I am learning </s> <s> I am happy I can study </s> "	
	V = (I,am,happy,learning,can,study)	
	V = (I,am,happy,learning,can,study, <unk>)</unk>	
	V = (I,am,happy,I,am)	
	V = (I,am,happy)	

In the context of our corpus, what is the estimated probability of word "can" following the word "I" using the bigram model and add-k-smoothing where k=3.

- $\bigcirc$  P(can|I) = 0
- O P(can|I) =1
- P(can|I) = 3/(2+3\*4)
- O P(can|I) = 3/(3\*4)
  - **⊘** Correct Correct.

O True				
False				
Correct.				



10. The perplexity score increases as we increase the number of <unk> tokens.</unk>	1 / 1 point
O True.	
False.	
Correct Incorrect.	