

## Feedback — Week 1 Practice Quiz

[Help Center](#)

Thank you. Your submission for this quiz was received.

You submitted this quiz on **Sun 14 Jun 2015 10:33 AM PDT**. You got a score of **12.00** out of **12.00**.

### Question 1

Which of the following word pairs are paradigmatically related? (select all that apply)

Your Answer		Score	Explanation
<input checked="" type="checkbox"/> computer, laptop	✓	0.25	
<input type="checkbox"/> car, drive	✓	0.25	
<input checked="" type="checkbox"/> car, vehicle	✓	0.25	
<input type="checkbox"/> computer, keyboard	✓	0.25	
Total		1.00 / 1.00	

#### Question Explanation

The correct answer is "car, vehicle" and "computer, laptop" since the word pairs tend to occur in the same context. We would generally not obtain a meaningful sentence by substituting "car" for "drive."

### Question 2

Which of the following word pairs are syntagmatically related? (select all that apply)

Your Answer		Score	Explanation
<input checked="" type="checkbox"/> car, drive	✓	0.25	
<input type="checkbox"/> car, vehicle	✓	0.25	
<input type="checkbox"/> computer, laptop	✓	0.25	

☒ computer, keyboard


0.25

Total

1.00 / 1.00

**Question Explanation**

The correct answer is "car, drive" and "computer, keyboard" since the words in each pair tend to co-occur in the same sentences.

## Question 3

Suppose the pseudo-document representations for the contexts of the terms A and B in the vector space model are given as follows:

 $d_A = (0.30, 0.20, 0.40, 0.05, 0.00, 0.05)$ 
 $d_B = (0.40, 0.10, 0.30, 0.00, 0.20, 0.00)$ 

What is the EOWC similarity score?

**Your Answer****Score****Explanation**
☐ 0.40

☐ 0.22

☐ 0.32

☒ 0.26


1.00

Total

1.00 / 1.00

**Question Explanation**
 $0.3 * 0.4 + 0.2 * 0.1 + 0.4 * 0.3 = 0.26$ 

## Question 4

EOWC treats words irrespective of their occurrence in the corpus. Adding IDF weighting to the EOWC similarity function will penalize common words in the corpus.

**Your Answer****Score****Explanation**
☒ True


1.00

☐ False

Total

1.00 / 1.00

## Question 5

The EOWC score is non-negative and cannot exceed 1.

**Your Answer**

**Score**

**Explanation**

☒ True



1.00

☐ False

Total

1.00 / 1.00

### Question Explanation

The EOWC score is the probability that two randomly selected words from the two pseudo-documents are the same. Since it is a probability, its values range between 0 and 1 (inclusive).

## Question 6

A random variable  $X$  with  $P(X=1)=1$  achieves the minimum possible entropy.

**Your Answer**

**Score**

**Explanation**

☒ True



1.00

☐ False

Total

1.00 / 1.00

### Question Explanation

This is a deterministic variable with  $H(X) = 0$ , which is the lowest possible value for entropy.

## Question 7

The outcome of an unbiased coin is easier to predict than the outcome of a biased coin.

## Your Answer

## Score

## Explanation

☐ True☒ False

1.00

Total

1.00 / 1.00

## Question 8

An unbiased coin has a higher entropy than any biased coin.

## Your Answer

## Score

## Explanation

☐ False☒ True

1.00

Total

1.00 / 1.00

### Question Explanation

An unbiased coin has  $H(X_{\text{unbiased}}) = 1$ , which is the maximum possible value of entropy for a binary random variable. Biased coins must have  $H(X_{\text{biased}}) < 1$ .

## Question 9

Which of the following is not true?

## Your Answer

## Score

## Explanation

☐ If  $H(X|Y) = H(Y|X)$  then  $H(X) = H(Y)$ ☐  $I(X;Y) = I(Y;X)$ ☒ If  $H(X) = H(Y)$  then  $X$  and  $Y$  follow the same distribution

1.00

Total

1.00 / 1.00

### Question Explanation

Counterexample for "If  $H(X) = H(Y)$  then  $X$  and  $Y$  follow the same distribution": Let  $P(X=1) = 0.9$  and  $P(Y=1) = 0.1$ . Clearly,  $H(X) = H(Y)$ , however, the distributions of the two random variables are different. "If  $H(X|Y) = H(Y|X)$  then  $H(X) = H(Y)$ " is always true because  $H(X) = I(X;Y) + H(X|Y)$ , and  $H(Y) = I(X;Y) + H(Y|X)$ .

## Question 10

Let  $X_{\text{text}}$ ,  $X_{\text{mining}}$ , and  $X_{\text{the}}$  be binary random variables associated with the words “text”, “mining”, and “the” respectively. Assume that the probabilities of the random variables are estimated based on a large corpus. Then we should expect  $H(X_{\text{text}}|X_{\text{mining}}) > H(X_{\text{text}}|X_{\text{the}})$ .

Your Answer	Score	Explanation
<input type="radio"/> True		
<input checked="" type="radio"/> False	✓ 1.00	
Total	1.00 / 1.00	

### Question Explanation

Since “mining” usually co-occurs frequently with “text”, it is expected to reduce the entropy of “text” more than “the” does.

## Question 11

$I(X;Y)=0$  if and only if  $X$  and  $Y$  are independent.

Your Answer	Score	Explanation
<input type="radio"/> False		
<input checked="" type="radio"/> True	✓ 1.00	
Total	1.00 / 1.00	

## Question 12

Let  $w$  be a word and  $X_w$  be a binary random variable that indicates whether  $w$  appears in a text document in the corpus. Assume that the probability  $P(X_w=1)$  is estimated by  $\text{Count}(w)/N$ , where  $\text{Count}(w)$  is the number of documents  $w$  appears in and  $N$  is the total number of documents in the corpus. You are given that “the” is a very frequent word that appears in 99% of the documents and that “photon” is a very rare word that occurs in 1% of the documents. Which word has a higher entropy?

**Your Answer****Score****Explanation**☒ Both words have the same entropy

1.00

☐ "the"☐ "photon"

Total

1.00 / 1.00

**Question Explanation**

Both words have the same entropy since  $H(X_{\text{the}}) = -0.99 \log(0.99) - 0.01 \log(0.01)$  and  $H(X_{\text{photon}}) = -0.01 \log(0.01) - 0.99 \log(0.99)$ .