## Congratulations! You passed!

**Grade received** 90% **To pass** 80% or higher

Go to next item

**1.** The minimum edit distance between the words *deep* and *creepy* is: **1 / 1 point**

4

**Correct**
That's correct. You need to replace *d* for *c*, which counts for 2, insert *r* and insert *y*.

**2.** Which of the following is a NOT VALID example of an edit string operation? **1 / 1 point**

- INSERT a letter: 'aple' --> 'apple'
- DELETE a letter: 'cloack' --> 'cloak'
- (selected) SWITCH a letter 'Lusca' --> 'Lucas'
- REPLACE a letter 'Crayom' --> 'Crayon'

**Correct**
Switching a letter is a valid operation ONLY when switching adjacent letters! In this case there were two switches: switch *s* and *c* and after *s* and *a*.

**3.** Autocorrect is only appliable when dealing with misspelled words. **1 / 1 point**

- (selected) False
- True

**Correct**
That's right, autocorrect can also be used for words that does not make any sense for a particular sentence. For instance, 'Happy birthday deer friends' is a correct spelled sentence, but the word *'deer'* makes no sense – it should be **dear**.

**4.** Given the corpus: **1 / 1 point**

"I am happy because I am doing quizzes."

Based on this tiny corpus, consider the following sentence:

"I **sm** very good at solving quizzes."

Which of the following is true?

- It is not possible to decide a correction for the misspelled word "sm".
- (selected) There is a unique correction for the misspelled word "sm".
- There is more than one possible candidate for a correction to the misspelled word "sm".
- The corpus is too tiny, so it is not possible to build a probabilistic model for autocorrection.

**Correct**
That's correct! The correction would be the word "am".

**5.** About the probabilistic model defined in the lecture, select all that apply. **1 / 1 point**

- [ ] Words with the same probability in the corpus will be equally likely to be candidates for a possible word correction.
- [x] Replacing a character costs more than deleting a character.

**Correct**
This is correct, replacing a word costs 2 whereas deleting it costs 1.

- [x] If $C(w)$ is the number of times a word appear in a corpus and $V$ is the corpus size, then the probability of the word $w$ in the corpus is $P(w) = \frac{C(w)}{V}$.

**Correct**
This is correct.

- [x] The sentence "Happy birthday deer friends" would not have any word corrected in the model defined in the lecture.

**Correct**
This is correct. Since the model just looks at misspelled words, the above sentence would not be corrected.

**6.** Suppose we build a distance matrix D for the following case: **0 / 1 point**

Source: Pie --> Target: Bye

What is the value for D[3,2]?

4

**Incorrect**
Please, revise the concepts introduced in the **Minimum edit distance algorithm** video.

**7.** About the Minimum edit distance algorithm, select all that apply. Let $D$ be the distance matrix, for two words of same size. The matrix size is $n$. **1 / 1 point**

- [x] $D[0, i] > D[0, j]$ if $i > j$.

  **Correct**
  This is correct, the first line will always have increasing values as we move to the right because it is the cost from editing the null string.

- [ ] $D[n, n]$ stores the highest value in the matrix.
- [ ] $D[i, j] = min(D[i-1, j] + \text{del\_cost}, D[i, j-1] + \text{ins\_cost}, D[i-1, j-1] + \text{rep\_cost})$
- [x] The algorithm avoids usage of brute force by implementing a dynamic programming approach.

  **Correct**
  That's correct. Using previous computed cells to compute another one is a dynamic programming method.

**8.** About the minimum edit distance, which of the following statement **is not true?** **1 / 1 point**

- ( ) It is used to evaluate similarity between two strings.
- (•) It is used to check if a word is misspelled.
- ( ) It counts the minimum number of edits to transform one string into another.
- ( ) It is used to implement spelling correction, document similarity and machine translation.

  **Correct**
  Correct! It is a measure between two strings and not a method to decide if a string is misspelled or not.

**9.** The minimum edit distance calculation is more computationally intensive if we have a big corpus. **1 / 1 point**

- ( ) True
- (•) False

  **Correct**
  That's correct. The minimum edit distance depends only on the editing cost and the two words that are being considered and not on any corpus or vocabulary.

**10.** Given the corpus "Autocorrect is a powerful tool and it is used on our computer." **1 / 1 point**

The value for $P(\text{is})$ is:

The answer should have two decimal places (rounding up, if necessary). For example: 0.88888 should be answered as 0.89.

0.17

**Correct**
That's correct!