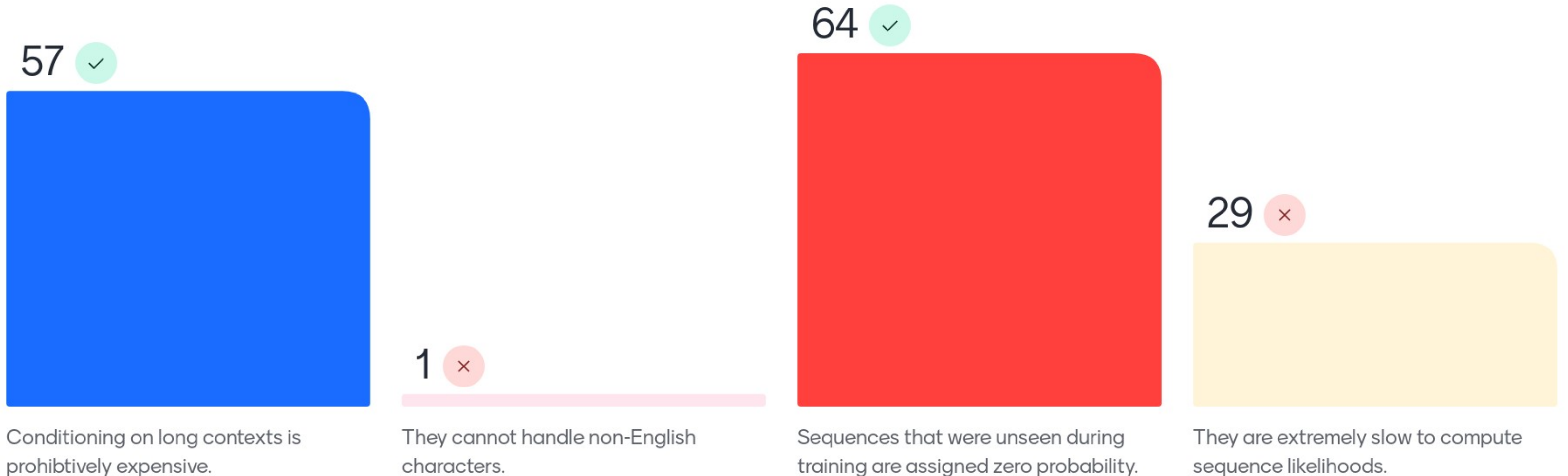
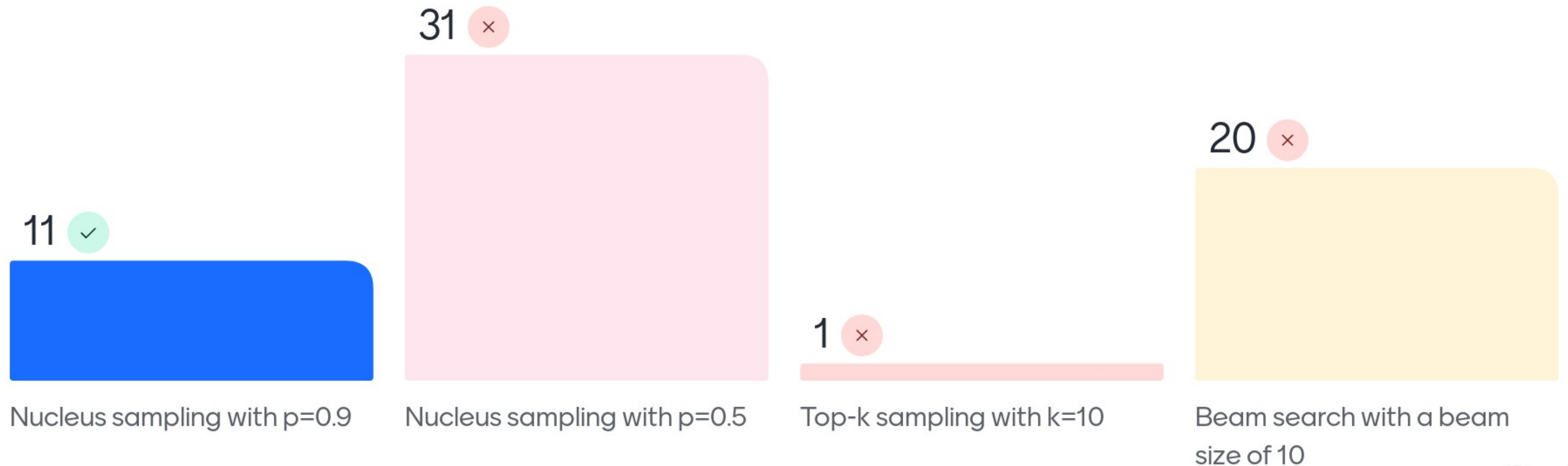


Which of the following are limitations of statistical language models that neural language models help solve?



Which of the following decoding strategies will result in the lowest likelihood generated text?



How do most subword tokenizers handle words that are not in their vocabulary?

62 ✓



Break the word down into substrings which are in the vocabulary

2 ✗



Break the word down into individual characters.

13 ✗



Replace the word with a special token, indicating it's unknown

0 ✗



Skip tokenizing this word