A2

The P+X approach involves removing the last word of each line, then using GPT-2 to predict the next word at a specified rank in its probability list. Changing the X value alters the predictability of the final output. When X is small, such as one or two, GPT-2 often chooses words that fit smoothly within the sentence, sometimes even matching the original text. This happens because GPT-2's highest-ranked options align closely with common language patterns and context. When X is higher, like 20, the substituted words tend to drift farther from the initial meaning. They can still form grammatical sentences, but they may sound odd or contain fragments such as "isn," which arises from GPT-2's subword tokenization process.

These variations in the chosen token reflect how GPT-2's probability distribution captures different linguistic possibilities. A higher X value often leads to more surprising or out-of-context word choices but still grammatically correct words. While a lower X value yields more predictable completions. It is interesting to see how consistent GPT-2 can be when predicting certain phrases or words at lower ranks, highlighting how patterns in text are learned and reproduced.

To implement a P+7 technique that specifically targets nouns, you could first detect which tokens in your input are nouns. You would feed GPT-2 enough context to predict a next token, but only accept noun candidates from GPT-2's predictions. You would then select the seventh most likely noun for each replacement. This method would ensure that only nouns are replaced, keeping other parts of speech untouched. The result would be a transformed sentence where each original noun is exchanged with GPT-2's seventh-ranked noun choice, leading to text that follows your desired positional probability constraint.