Homework Assignment No. 1

Title: Meaning and Nonsense: The Library of Babel

Part 1

P1Q1: What does the Library of Babel say about the nature of language and meaning?

Borges' work suggests that language and meaning are not a numbers game. In containing "all possible combinations of the twenty-two orthographic symbols" the library contained far more nonsense than rational thought. For language to have meaning, language requires the human touch, which brings structure, intention, and context.

P1Q2: How is this story related to how NLP systems process or generate language?

Like Borges' Library, NLP systems generate text systematically, producing outputs that can appear to be filled with intention, meaning, and insight. They can also take the form of pure hallucination. Just as the librarians cannot distinguish meaningful books from plausible falsehoods among infinite combinations, users of language models face the challenge of identifying accurate information within statistically probable but potentially fabricated responses.

P1Q3: How might the text from the Library of Babel be useful for NLP?

The text from the Library of Babel could be useful for NLP as a training dataset to help models distinguish between meaningful and meaningless text, teaching them to recognize patterns that indicate coherence versus random character combinations (i.e. M C V). It could also serve as a benchmark for evaluating how well language models can identify and filter out nonsensical outputs, similar to how the librarians struggled to find meaningful books among the infinite gibberish.

Part 2 Meaning in the making

Samples are ranked in the table below:

Rank	Sample	Texts
1	1	"The sun, weary from its long descent, brushed the edge of the horizon with gold. In the silence that followed, only the trees remembered the day."
2	2	"Frondulous gleams converge in lexical harbors where syntax is sublimated by infinite recursion."
4	3	"The the of into and not he had it as was on in by."
3	4	"A kind in glass and a cousin, a spectacle and nothing strange a single hurt color and an arrangement in a system to pointing. All this and not ordinary, not unordered in not resembling. The difference is spreading."
5	5	"9zxtrplmnn qvek wj-289, jkqwe qux mvxz."

I ranked these texts by how much cognitive wiggle room I had to allow for a statement to make sense – if any. For instance sample 1 requires no wiggle room, but sample 2 appears like it might read as a sentence if the first work were "Fabulous" instead of "Frondulous". Sample 4 reads like it could be a lyric out of a Bob Dylan song. Samples 3 and 5 are beyond any amount of cognitive wiggle room as they make no sense.