

Tarea de Seminario I

Realizar un programa en ANSI C el cual tome como entrada un archivo de texto plano(en este caso 'cuento.txt') y como salida escriba en otro archivo la matriz de las probabilidades de unigrama, bigrama y trigramas de cada palabra que esté contenida en el texto.

Por lo tanto el archivo de salida debe contener algo como esto:

<i>Palabra / Palabra(s) Anteriores</i>	\emptyset	así	es	donde	es, eso	es, así
como	0.3	0.1	0,15	0.02
espejo	0.05	0	0.001
tierra	0.1	0
mujer	0,09	0
...
...

Recuerde:

- Unigrama: $P(w_n)$
- Bigrama: $P(w_n | w_{n-1})$
- Trigramas: $P(w_n | w_{n-1}, w_{n-2})$