

Resultados cuantitativos del buscador proximal

Tabla de precisiones para las colecciones *Clueweb-1K* y *Clueweb-10K*.

		Clueweb-1K		Clueweb-10K	
		P@5	P@10	P@5	P@10
ProximalSearcher	Basic Index	0.44	0.46	0.20	0.17
	Stopwords Index	0.44	0.44	0.20	0.19
	Stem Index	0.32	0.42	0.12	0.15
	Advanced Index	0.36	0.44	0.19	0.16

Para obtener estos valores se ha utilizado la clase **SearcherTest** implementada en la práctica anterior sobre las dos colecciones que es capaz de indexar nuestra aplicación.

Ejemplo de uso del buscador proximal

Se ha comprobado el correcto funcionamiento de la implementación de ProximalSearcher haciendo uso de los ejemplos suministrados en las diapositivas de teoría.

Ejemplo

And by opposing end them: **to** die, **to sleep**

No more; and by a **sleep, to** say we end

The Heart-ache, and the thousand Natural shocks

That Flesh is heir **to**? 'Tis a consummation

Devoutly **to** be wished. **To** die **to sleep,**

To sleep, perchance **to** Dream; Aye, there's the rub,

For in that **sleep** of death, what dreams may come,

When we have shuffled off this mortal coil,

Must give us pause.

$q = \text{to sleep}$

$score(d, q) = \frac{1}{1} + \frac{1}{1} + \frac{1}{1}$

$+ \frac{1}{1} + \frac{1}{1} + \frac{1}{2} + \frac{1}{10} = 5.6$

a	b
$-\infty$	
7	8
14	15
41	42
42	43
43	44
44	46
46	56
	∞

En este caso, definiendo los Posting de cada uno de los términos, la salida del programa es la siguiente:

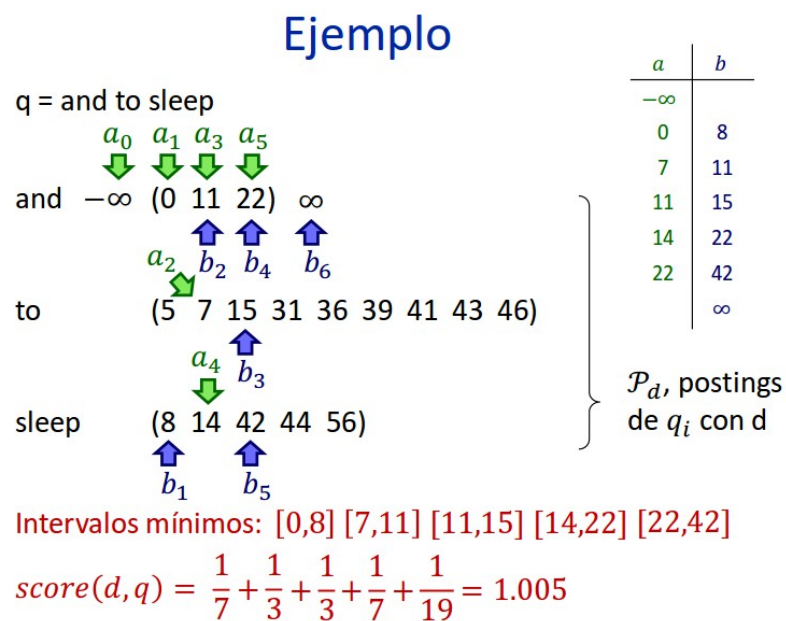
```

1 [7,8]
2 denominator = 1.0
3 [14,15]
4 denominator = 1.0
5 [41,42]
6 denominator = 1.0
7 [42,43]
8 denominator = 1.0
9 [43,44]
10 denominator = 1.0
11 [44,46]
12 denominator = 2.0
13 [46,56]
14 denominator = 10.0
15 score = 5.6
16 |

```

Denominator se refiere al valor que toma el denominador a la hora de realizar el cálculo del score ($b - a - |q| + 2$).

En el otro ejemplo:



81

La salida del programa es la esperada:

```

1 [0,8]
2 denominator = 7.0
3 [7,11]
4 denominator = 3.0
5 [11,15]
6 denominator = 3.0
7 [14,22]
8 denominator = 7.0
9 [22,42]
10 denominator = 19.0
11 score = 1.0050125313283207

```