

Compiladores

“Le hacemos al léxico”

Dragon Slayers

8 de octubre de 2018

El rendimiento del análisis léxico es vital. Compararemos varias opciones para efectuar una tarea típica: aislar «palabras» en archivos grandes (similar a lo que hace el programa `wc(1)`). Para nuestros efectos, una palabra es una secuencia de letras, posiblemente unidas por guión (como son `letrado` o `no-terminal`; note que por ejemplo `O'Higgins` cuenta como dos palabras, `Entel123` o `3Com` cuentan como una y `1234` o `3,141592` no cuentan). Para simplificar, considere texto puramente ASCII (texto internacional tiene sus propias enternecedoras complicaciones...). Use las siguientes alternativas:

1. Escrito a mano, leyendo caracter a caracter vía `getc(3)`.
2. Escrito a mano, leyendo el archivo completo a memoria con `read(2)` y procesando allí
3. Escrito a mano, mapeando el archivo completo a memoria con `mmap(2)` y procesando allí
4. Usando `flex(1)`, un *scanner* interactivo
5. Usando `flex(1)`, un *scanner* no interactivo

Resuma sus resultados y conclusiones.

1. Condiciones de entrega

- La tarea se realizará *individualmente* (esto es grupos de una persona), sin excepciones.
- La entrega debe realizarse vía [Moodle](#) en un *tarball* en el área designada al efecto, bajo el formato `tarea-1-rol.tar.gz` (`rol` con dígito verificador y sin guión).

Dicho *tarball* debe contener las fuentes en LaTeX (al menos `tarea.tex`) de la parte escrita de su entrega, además de un archivo `tarea-1.pdf`, correspondiente a la compilación de esas fuentes.

- En caso de haber programas, su ejecutable *debe* llamarse `tarea-1`, de haber varias preguntas solicitando programas, estos deben llamarse `tarea-1-1`, `tarea-1-2`, etc. Si hay programas compilados, incluya una `Makefile` que efectúe las compilaciones correspondientes.

Los programas se evalúan según que tan claros (bien escritos) son, si se compilan y ejecutan sin errores o advertencias según corresponda. Parte del puntaje es por ejecución correcta con casos de prueba. Si el programa no se ciñe a los requerimientos de entrada y salida, la nota respectiva es cero.

- Además de esto, la parte escrita de la tarea debe en hojas de tamaño carta en Secretaría Docente de Informática (Piso 1, edificio F3).

- Tanto el *tarball* como la entrega física deben realizarse el día indicado en [Moodle](#). No entregar la parte escrita en papel o no entregar en formato electrónico tiene un descuento de 50 puntos.

Por cada día de atraso se descontarán 20 puntos. A partir del tercer día de atraso no se reciben más tareas, y la nota de la tarea es cero.

- Nos reservamos el derecho de llamar a interrogación sobre algunas de las tareas entregadas. En tal caso, la nota base (antes de descuentos por atraso y otros) es la de la interrogación. No presentarse a la interrogación sin justificación previa significa automáticamente nota cero.