



4º Ingeniería Informática

II26 Procesadores de lenguaje

Control sobre el léxico, grupo TE-1, 14 de noviembre de 2008 (solución)

PREGUNTA 1

(0,5 PUNTOS)

Para cada uno de los siguientes lenguajes, escribe una expresión regular que lo represente o un AFD que lo reconozca (no es necesario que escribas ambos).

1. Secuencias de dos o más letras vocales en las que la última letra no está repetida. Por ejemplo, son válidas las cadenas *aeao*, *iiiu* y *aeiou* pero no *aeae* ni *oeoe*.

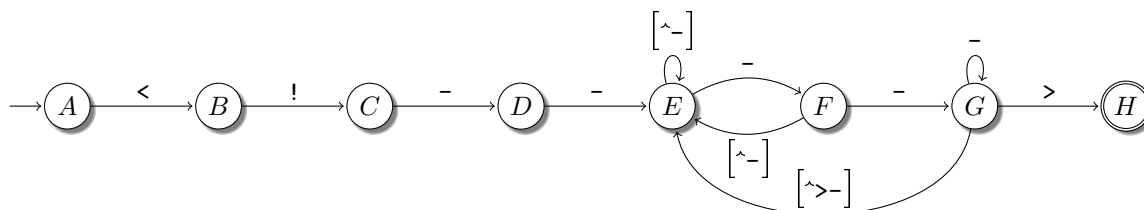
La expresión regular siguiente serviría si sólo consideramos las vocales minúsculas:

$$[aeio]^+u|[aeiu]^+o|[aeou]^+i|[aiou]^+e|[eiou]^+a$$

La generalización al caso de mayúsculas y minúsculas es trivial. El AFD resulta bastante complicado (si aplicamos el algoritmo de transformación, necesitamos 57 estados).

2. Comentarios que comienzan por la secuencia `<!--` y terminan con la primera aparición de `-->` que no se solape con ella. Por ejemplo, son válidos los comentarios `<!-- esta es -mi- página -->` o `<!-- f(a)->b -->` pero no `<!-->` o `<!-- a->`. Estos comentarios son similares, pero no iguales a los de HTML.

En este caso, la solución más sencilla es el AFD:



Cambián es posible escribir expresiones regulares. Por ejemplo, las siguientes:

- `<!--(-*[^>-]|-?>)*-+>`
- `<!--(-?[^-]|-*[^->])*-+>`
- `<!--([^-]*-([^-]|-+[^>-]))*[^-]*-+>`

Duración del control: 1½ horas.

¡Buena suerte!