

SHAKA.cnv

Gutiérrez Cristóbal Ignacio
Facultad de Ciencias de la Computacion
Benemerita Universidad Autonoma de Puebla
14 Sur y Av San Claudio, Ciudad Universitaria 72570, Puebla, México
E-mail: tachouhoner@hotmail.com

KEYWORDS

Expresiones regulares, Convert, REC, XML, XLS.

Figures 2: Contenido de la Etiqueta

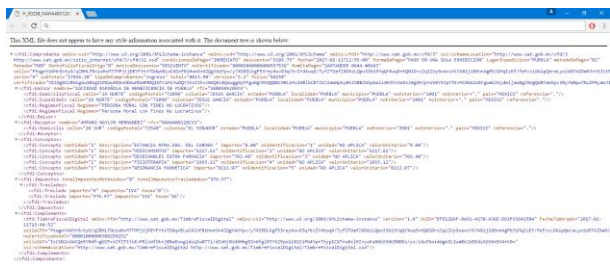
RESUMEN

Este es el problema presentado por una dependencia gubernamental que intenta extraer cadenas de texto de varios archivos y agregarlas a otro archivo. Daremos solución usando expresiones regulares bajo el lenguaje de programación Convert y su Compilador REC, aportaciones del Dr Harold V. McIntosh.

Siguiendo un orden, adelantaremos los pasos a realizarse inmediatamente:

- ❖ Leer datos
- ❖ Procesar datos
- ❖ Entregar resultado

INTRODUCCION



Figures 1: Contenido del Archivo

Los archivos de donde se extraerán las cadenas de texto, son de tipo XML. Al ser archivos de etiquetado es fácil encontrar las cadenas buscadas. La etiqueta a buscar será “UUID” y su contenido es el que se procesara. En los requerimientos se solicita que el archivo generado sea uno de tipo Excel.

ANTES DE EMPEZAR

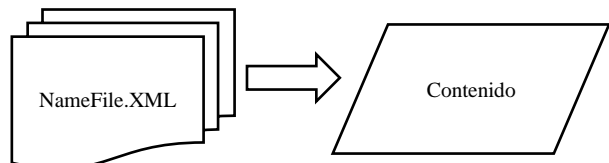
Al utilizar Conver y REC hay que realizar ciertos ajustes a los Nombres de los Archivos a utilizar para el buen funcionamiento.

Estos lenguajes permiten usar archivos con Maximo 7 caracteres en Nombre y 3 caracteres en Extension. El editar el nombre del archivo no afecta al contenido de archivo.

LECTURA DE DATOS

Primero leeremos todo el contenido del archivo para poder procesarlo debidamente.

Mientras el archivo sea de extensión “XML” leera el contenido hasta encontrar el carácter de fin de archivo “EOF”.

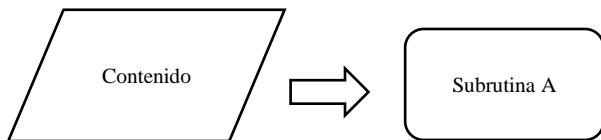


Figures 3: Proceso de Lectura



PROCESAMIENTO DE DATOS

El procesamiento se hara desde una subrutina, la cual trabajara en base a expresiones regulares. Se mandara todo el contenido del archivo a esta subrutina, que buscara la etiqueta “UUID” y cuando la encuentre, devolverá su contenido que se delimita por <<“>...<”> (Comillas).



Figures 4: Procesamiento de los Datos

“Convert es un lenguaje dirigido por patrones, cuyos mandatos se escriben como paradigmas en el estilo tipificado por la frase: ‘Si ves esto, Haz esto otro’ ” (McIntosh, 1986)

La siguiente línea muestra el trabajo de la Subrutina A:

<1>UUID=“<2>”<3> (1)

Haciendo una analogía con listas, recordemos que, todas están compuestas por una Cabeza (Head) de lista y el Resto (Rest) de la lista.

<1> Es la variable que Casara (Match) con la Cabeza del texto.

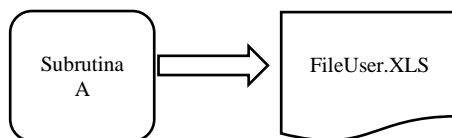
UUID=“ Es la cadena a buscar en el documento.

<2> Es la variable que cazara con el contenido que deseamos.

” Es el carácter que debe encontrar para determinar donde acaba la cadena deseada.

<3> Es la variable que casa con el resto del contenido.

La cadena buscada ahora se encuentra almacenada en la variable <2>. El contenido de esa variable se guarda en el archivo que nos indique el usuario. Si el archivo no existe, crea uno nuevo. Si ya existe, el contenido se almacenara al final del archivo.



Figures 5: Guardar los Datos

vimos la oportunidad para usar el Compilador de Expresiones Regulares REC, la Magnum Opus de aquel hombre que decidió quedarse en medio de la nada solo para enseñarnos a hacer Ciencia. Ave Mc, los que vamos a programar te saludamos.

REFERENCES

Notas sobre los Lenguajes REC y Convert
Gerardo Cisneros y Harold V. McIntosh
ICUAP, Puebla, México. 1986.

La computación y los procesos iterativos
Comentarios del Dr. Harold V. McIntosh
Publicado por Alberto Licon A., Victoria R. Bautista y M.A. Burgos

REC- A regular expression Compiler
Gerardo Cisneros and Harold V. McIntosh

Tesis profesional Compilador de Convert para E!1
procesador Intel 8086
Jesús Ariel Cari-asco Ochoa.

Tesis Profesional Biblioteca de Convert para
Procesador Intel 8086
Lucía Angélica cae la Vega Doria.

ENTREGA RESULTADO

Al terminar de procesar todos los archivos con extensión “XML” ubicados en ese directorio, existirá un archivo de Excel con las cadenas de cada uno de los archivos, uno en cada fila, disponible para trabajar, ver y editar su contenido.

CONCLUSION

Al solucionar este pequeño problema esperamos grandes cambios. La solución no es única y se puede realizar en el Lenguaje de Programacion que se desee. Al ser un problema que se resuelve con Expresiones Regulares,