

REPRESENTACIÓN DE SERIES CON PATRONES ANIDADOS

GERARDO L. FEBRES

RESUMEN

El análisis clásico de procesos unidimensionales periódicos se soporta en la aplicación de las Series de Fourier y el algoritmo FFT. En contraste, la caracterización de procesos no-periódicos es difícil por lo escaso de herramientas efectivas. Este estudio propone métodos para la creación de espacios en los cuales se pueda detectar y cuantificar el peso relativo de patrones. En ese sentido se proponen algoritmos para la detección de secuencias de caracteres repetidas en textos. La representación de textos como la concatenación de frases repetidas de diversa longitud permite la caracterización de textos en el plano longitud de secuencia vs frecuencia. Cuando el proceso viene descrito como una secuencia de valores numéricos, la conversión de la serie a texto se hace con la aplicación de escalas para la discretización de los valores. Para la detección de patrones, definidos aquí como secuencias de caracteres que aparecen un número especificado de veces o más, se elaboró un conjunto de algoritmos que manejan en un tiempo razonable, textos de unos 40000 caracteres. Los algoritmos funcionan en el ambiente de trabajo *MoNet*: una plataforma para el modelado de experimentos complejos que desde 2011 ha desarrollado el prof. Gerardo Febres. Se presentan experimentos realizados con series de tiempo asociadas a la Ecuación Logística de May, Ecuaciones de Lorentz y valores del mercado bursátil. Más allá de los experimentos presentados en este trabajo, la aplicación del método ofrece utilidad en la caracterización de series de tiempo no necesariamente periódicas, así como estructuras de otra índole como textos literarios, música, y cadenas de ADN. Algunas consecuencias de este desarrollo es la posibilidad de evaluar la complejidad multi-escala de sistemas a través de la cuantificación de la entropía a varias escalas de observación. Incluso, la superposición de secuencias anidadas se manifiesta como una disminución de los valores de entropía calculados que, circunstancialmente, puede alcanzar valores negativos. Es posible que los balances de información que pueden construirse con estos valores de entropía, sean una oportunidad para cuantificar la redundancia en textos; un viejo reto dentro de la Teoría de la Información.

Palabras clave: **Patterns**; nested patterns; information; entropy; ordered structure; redundancy

UNIVERSIDAD SIMÓN BOLÍVAR: *Email address*: gerardofebres@usb.ve