

LE, CUADERNO DE LECTURA COMPARADA. UTILIZACIÓN DEL TEXTO GENÉTICO Y EL GRAMATICAL A TRAVÉS DEL ALINEAMIENTO DE SECUENCIAS.

ELOI PUIG MESTRES.

Facultad de Bellas Artes de Barcelona, Departamento de Pintura, Universidad de Barcelona.
puigme@ub.edu

Resumen

Este artículo pretende presentar una lectura performática entre dos actores-lectores en directo de un fragmento del cuaderno *LE* a modo de *Alfabeto por palabras (Spelling Alphabet)* que, a través del método de *Alineación de secuencias* compara dos textos, uno literario y otro científico. Para esta lectura se ha creado *Bocal*, un software que interpreta las palabras y permite construir la frase previamente seleccionada y proyectarla al público.

LE forma parte de la *Colección Alineamientos*, configurada por una serie de cuadernos que comparan textos científicos y literarios. Se trata de una visualización creada por Aleix Molet, doctor en Bellas Artes y Eloi Puig, artista e investigador de la Universidad de Barcelona, en torno a la relación existente entre el código gramatical de un texto y la codificación genética de las cadenas de ADN. Esta experiencia fue desarrollada gracias a la colaboración del equipo de Genómica Computacional (Computational Genomics | BSC-CNS)¹ dirigido por David Torrents.

Un *Alineamiento de Secuencias* en bioinformática es una forma de representar y comparar dos o más secuencias de ADN para resaltar zonas de semejanza que indican relaciones funcionales o evolutivas entre los genes. Nuestro objetivo fue modificar este método para poder comparar texto gramatical.

Palabras-clave: METAMÉTODO, INVESTIGACIÓN ARTÍSTICA, ALINEAMIENTO DE SECUENCIAS, ARTE Y CIENCIA, ALFABETO POR PALABRAS, PERFORMANCE.

Abstract

This paper aims to present a reading performance of a piece of the notebook *LE* as a *Spelling Alphabet* between two actors-readers. *LE* compares two texts, one literary and one scientist through the *Sequence Alignment* method. For this reading has created *Bocal*, software program that interprets the words and allows to build the previously selected sentence and projects to the public.

LE is part of the *Alignments Collection*, set a series notebooks comparing scientific and literary texts. This is a visualization of two artists from the University of Barcelona about the relationship between a grammatical text and DNA code. This experience was developed through collaboration Computational Genomics team led by David Torrents from BSC at the Polytechnic University of Barcelona.

A *Sequence Alignment* in bioinformatics is a way to represent and compare two or more DNA sequences to highlight areas of similarity indicating functional or evolutionary relationships between genes. Our goal was to modify this method to compare grammatical text.

Keywords: METAMETHOD, ARTISTIC RESEARCH, SEQUENCE ALIGNMENT, ART AND SCIENCE, SPELLING ALPHABET, PERFORMANCE

1. INTRODUCCIÓN

Ambos artistas, Aleix Molet y Eloi Puig, del equipo de investigación artística Imarte de la facultad de Bellas Artes de Barcelona, hemos estado trabajando con el equipo de Genómica Computacional dirigido por David Torrents como actividad del proyecto de investigación I+D Metamétodos². Este proyecto se centra en nuevas estrategias metodológicas de visualización e interpretación de los procesos del arte. Busca conectar con las estructuras de investigación que otras disciplinas aplican en sus proyectos, para de esta manera, poder entenderlas, leerlas y posteriormente, en una segunda etapa, poder adaptarlas y modificarlas ya desde una perspectiva creativa, permitiendo que surja un pensamiento original, nuevas ideas y nuevos conceptos. Se trata, en definitiva, de un metamétodo que trata (desde el arte como proceso y producción) encontrar nuevas relaciones entre la teoría y la práctica, entre la ciencia y la tecnología creando una serie de enlaces multidisciplinares basados en la investigación, estableciendo vínculos de similitud y diferencia entre el método científico y el artístico.

La relación con el equipo de Genómica Computacional se inició en septiembre del año 2011 y se realizaron 7 diferentes encuentros hasta junio de 2012. El objetivo principal fue conocer qué tipo de investigaciones se estaban desarrollando y qué métodos utilizaban. Construimos un blog donde transcribir cada uno de los encuentros, poniendo de relieve la importancia del proceso para llegar a detectar el método que utilizaríamos para dar respuesta a lo que, como artistas, observábamos y entendíamos.

El diálogo fue muy fructuoso, ya que entendimos la relevancia del *Alineamiento de Secuencias* como principio del trabajo desarrollado por el grupo de investigadores del equipo de Genómica. Este método de alineamiento fue modificado para poder comparar, no cadenas de ADN, sino texto gramatical, aquellos conjuntos de caracteres que siguen un conjunto de reglas y principios que gobiernan el uso de una lengua concreta determinada.

De esta manera una vez verificado y acordado una línea posible de convergencia entre los investigadores de genómica y nosotros, como intérpretes del ámbito artístico, iniciamos una serie de experimentos basados en el alineamiento, que dieron lugar un conjunto muy variado de propuestas. Todas ellas han sido mostradas en diferentes espacios expositivos y bajo múltiples formatos, como por ejemplo:

1. *Sara sings*. Alineación entre la letra de una canción infantil y la transcripción del canto de la misma canción hecha por una niña de 3 años. Sala Corretger5. Barcelona. 2012.
2. *TE LA (Text Local Alignment)*. Alineación entre la transcripción de dos vídeos domésticos sobre la relación entre ciudad y ciencia. Arts Santa Mónica. Barcelona. 2012.
3. *Renga-Ronda*. Alineación entre el poema de un poeta escocés y el poema de un poeta catalán. HICA, Escocia. 2012.
4. *The White Wall*. Lectura de una frase de J. Guillén de *El argumento de la obra*, sobre el espacio de lectura en el que se encuentra el texto escrito. Galería F. Cano. 2013.
5. *Wagensber's Alphabet*. Alineación de aforismos de Jorge Wagensber. Fresados planográficos sobre planchas blancas de PVC. Julio 2014.
6. HM. Interpretación de un fragmento del cuaderno HM a través del software *roloc* (traducción de imágenes en franjas cromáticas). Sala de exposiciones *El Carme*. Badalona 2014.

2. UN PARTICULAR USO DEL METODO DE ALINEMAMIENTO DE SECUENCIAS

El ser humano se origina con una sola célula formada a partir de un óvulo y un espermatozoide, pero para que esta célula única se convierta en los miles de millones de células que forman el cuerpo humano, y toda esta máquina celular funcione de una forma coordinada, hay un código con instrucciones, el ADN. El código o texto del ADN se basa en el código ATCG y, dependiendo de los intereses de la investigación genética, la cantidad de caracteres puede variar. Aunque los software de alineamiento que se usan comúnmente tienen únicamente el código de 4 caracteres A-T-C-G, los genetistas que han desarrollado el software Swat5³, para ser más precisos han asignado a cada tipo de par de bases una letra del abecedario, en total 28. Existen 4 pares de bases⁴, y estas interactúan de diferentes maneras unas con otras mediante enlaces llamados *puentes de hidrógeno*. De esta forma, según el tipo de base y los diferentes enlaces que puede formar existirían 28 tipos de pares de bases. En el caso que aquí presentamos, la investigación en bioinformática hace necesarias un número de letras igual al número de letras del alfabeto. Ambos códigos, el genético y el gramático, utilizan 28 caracteres.

Con el objetivo de detectar cambios genéticos entre cadenas de ADN, el *Alineamiento de Secuencias* es el procedimiento básico de análisis por comparación. Es básico pero no sencillo, ya que este método trata de dar solución a los exigentes cálculos que se derivan de los estudios genéticos. Existe una gran variedad de software que ofrece diferentes formas de alineamiento, se diferencian en dos grandes grupos, los alineamientos locales o los globales. Por lo general, un alineamiento selecciona la parte más coincidente entre ambas líneas de código, marca en mayúsculas las coincidencias totales y las coincidencias aproximadas en minúsculas.

“Los alineamientos se representan normalmente con un formato gráfico y de texto. [...] Muchos programas de visualización de secuencias utilizan también esquemas coloreados para mostrar información de las propiedades de los elementos secuencia individuales”⁵.

Utilizar el *Alineamiento de Secuencias* supuso establecer el punto de partida de nuestra investigación. Los códigos de la bioinformática y la gramática de los textos utilizan el mismo número de caracteres. Esa coincidencia nos sirvió para sustituir los 28 caracteres del código genético por los 28 caracteres de nuestro alfabeto. Utilizando el mismo software bioinformático alineamos texto gramático, literario. Swat5 nos fue facilitado por los investigadores genómicos del BSC y, al mismo tiempo, se nos introdujo a su utilización. Los resultados de estas alineaciones son mostradas con el mismo aspecto gráfico y formal con el que los científicos realizan sus estudios, estadísticas y conclusiones. En nuestras alineaciones aparecen en mayúsculas las coincidencias totales y las coincidencias aproximadas en minúsculas, espacios en blanco y guiones negros para ajustarlas. Serán alineamientos teniendo en cuenta que las proteínas de las cadenas de ADN se representan utilizando un código basado en 28 caracteres de texto genético, nuestro procedimiento sustituye estas cadenas de proteínas de ADN por texto gramático.

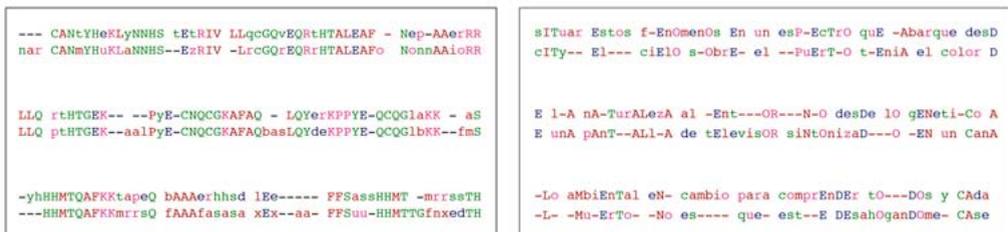


Fig.1. Ejemplos de alineaciones a través del mismo software Swat5. Imagen izquierda: Alineación de códigos de ADN. Imagen derecha: Alineación de textos gramaticales.

3. COLECCIÓN ALINEAMIENTOS

La *colección Alineamientos* se presenta en forma de cuaderno y consta de 4 números autoeditados. Cada publicación está formada por la alineación de un fragmento de dos libros, cada uno de ellos es seleccionado por un científico y por alguien del ámbito de las humanidades, conformando parejas que muestran sus propios e individuales intereses.

Este proyecto quiere evaluar la relación entre las dos culturas, la ciencia y las humanidades, utilizando el método del *Alineamiento de Secuencias* para visualizar aproximaciones al concepto de *Tercera Cultura* (Brockman 2000). *La Tercera Cultura* como la superación de las dos culturas: la científica y la humanística, para integrarlas en una tercera en la que sólo hay conocimiento.

El procedimiento para llevar a cabo el alineamiento sigue siempre una misma pauta. En primer lugar para cada número hemos conectado con dos agentes diferentes, uno relacionado profesionalmente con el ámbito científico y otro con el ámbito humanístico y/o artístico. A cada uno de ellos se les ha preguntado por una de las lecturas que han considerado importantes o que han detectado una especial afinidad y, posteriormente, nos han comunicado su referencia bibliográfica. Una vez que obtenemos el texto o libro referenciado, hemos seleccionado un fragmento para alinearlos mediante *Swat5*.

Los resultados de estos alineamientos han sido formalizados mediante una publicación. Hasta el momento tenemos los siguientes 4 números:

- no1 --- QN *¿Qué es lo que no sabemos?* de I. Prigogine y *Nichi nichu kore konichi* de J. Cage seleccionado por A. Molet y E. Puig.
- no2 --- HM *Qué es lo que nos hace humanos* de M. Ridley y *Neuromancer* de W. Gibson seleccionados por D. Torrents y D. Casacuberta.
- no3 --- LE *Ecología* de R. Margalef y *El Danubio* de C. Magris seleccionado por S. González y V. Magalhaes.
- no4 --- IJ *Chris Crawford On Interactive Storytelling* de C. Crawford y *El jardín de senderos que se bifurcan* de J. L. Borges seleccionado por A. Galbete y I. Pàmies.



Fig. 2. Los cuatro números de la Colección *Alineamientos*. Año 2014 y 2015

3.1. LECTURA DE UN FRAGMENTO DEL CUADERNO Nº3 LE

En este cuaderno se han comparado (alineado) dos textos de diferentes naturalezas. En primer lugar tenemos el libro seleccionado por el investigador en biología, Santiago González que escogió *Ecología* del naturalista y ecólogo español R. Margalef, pionero en la introducción de los estudios de ecología y en el desarrollo de dicha disciplina durante el s. XX. Por otro lado tenemos la novela *El Danubio* seleccionada por V. Magalhaes, artista y profesor de la Universidad de Madeira y escrita por el italiano C. Magris.

¿Qué comparten ambos textos? ¿qué nos aportan esas similitudes morfológicas entre ambos textos? Estas preguntas generan la performance formalizada a través de dos Lectores.

Cada Lector representa a uno de los textos y, en forma de confrontación, discusión, van generando la alineación en directo, frente al público, en la sala de exposiciones.

Para generar la alineación en directo se ha creado una aplicación específica a la que hemos nombrado *Bocal*. Su función simplemente es la de mostrar el alineamiento y la particularidad que la hace singular es el procedimiento en como lo hace. *Bocal* se activa por la voz y solo traduce palabras del *Alfabeto por palabras (Spelling Alphabet)* a su correspondiente vocal (Alpha → A, Bravo → B, etc...). La duración de esta lectura performática será aproximadamente entre 10 y 12 minutos.

Referencias

Brockman, John (ed.). 2000. *La Tercera Cultura. Más allá de la revolución científica*. Barcelona: Tusquets, Matemas.

Notas

- 1 Computational Genomics | BSC-CNS. [página web] <http://www.bsc.es/life-sciences/computational-genomics>. [accedido 02_06_2015] El Barcelona Supercomputing Center–Centro Nacional de Supercomputación es el centro pionero de la supercomputación en España. Su naturaleza es doble: por una parte es un centro de investigación formado por más de trescientos científicos, y por otra es un centro de servicios de supercomputación para toda la comunidad científica. También gestiona la Red Española de Supercomputación (RES). <http://bsc.es>
- 2 *Metamétodo: Metodologías compartidas y procesos artísticos en la sociedad del conocimiento* HAR2010-18453 (subprograma ARTE). Plan Nacional de I+D+i 2008-2011. Subprograma de proyectos de investigación Fundamental no orientada. Subvencionado por el Ministerio de Economía y Competitividad y la Unión Europea.
- 3 Swat5 fue escrito en el año 2010 por Friman Sanchez del Departamento de Arquitectura Computacional de la Universidad Politécnica de Catalunya.
- 4 En genética un par de bases consiste en dos nucleótidos opuestos y complementarios en las cadenas de ADN y ARN que están conectadas por puentes de hidrógeno. En el ADN adenina y timina así como guanina y citosina, pueden formar un par de bases. En ARN, la timina es reemplazada por el uracilo conectándose este con la adenina. https://es.wikipedia.org/wiki/Par_de_bases [accedido 14_05_2015]
http://es.wikipedia.org/wiki/Alineamiento_de_secuencias [accedido 14_05_2015]