GLOBAL COHERENCE, LOCAL UNCERTAINTY

– A RECIPE FOR LITERARY SUCCESS

Kristoffer Nielbo, Yuri Bizzoni, Ida Marie S. Lassen,
Pascale Feldkamp Moreira, Mads Rosendahl Thomsen,
Jianbo Gao

AARHUS
UNIVERSITY
CENTER FOR HUMANITIES COMPUTING

CENTER FOR
HUMANITIES
COMPUTING

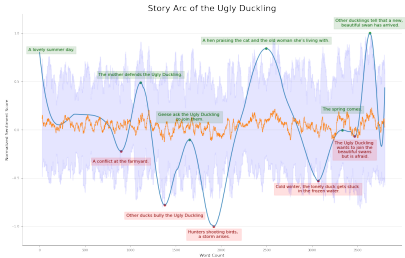
OUTLINE

1. Introduction
2. Methods
3. Results
4. Discussion

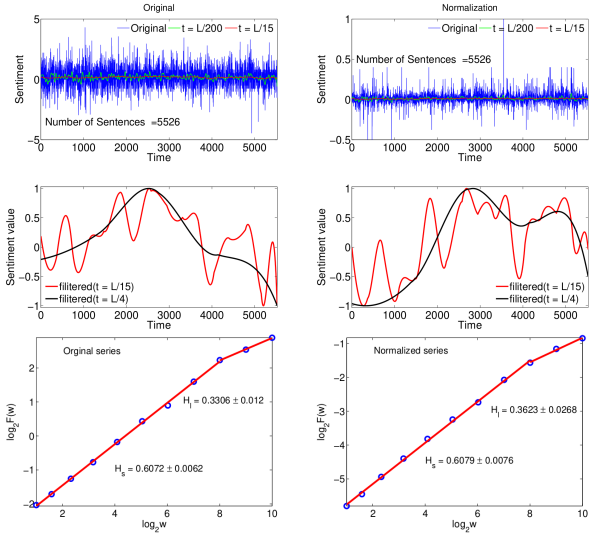
OUTLINE

A Recipe for Literary Success

- **context dependency** – 'the success of a work of literature depends entirely on its context'
- **work internalism** – 'success depends on work-internal features'
- perception and transmission of literature facilitates the *emergence of specific properties of successful literature*



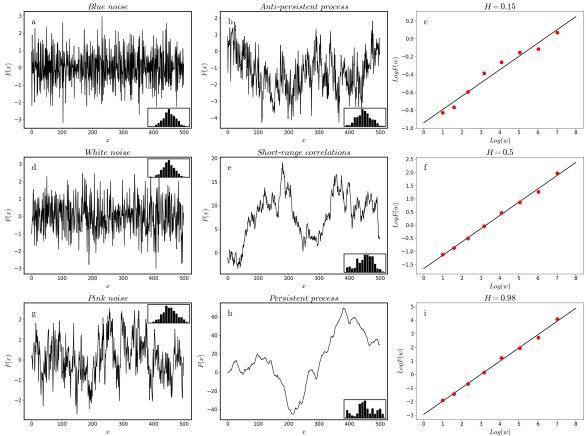
Story arc of the Ugly Duckling by H.C. Andersen.



The narrative arc for noble laurate Kazuo Ishiguro's *Never Let Me Go* - an example of a successful novel [1].

OUTLINE

1. Introduction
2. **Methods**
3. Results
4. Discussion



By computing the global fits, the residual, and the variance between original random walk process and the fitted trend for each window size w, we plot $\log_2 F(w)$ as a function of $\log_2 w$. The presence of fractal scaling amounts to a linear relation in the plot, with the slope of the relation providing an estimate of H.

Detrended fluctuation analysis

Peng et al 1994 method for estimation of H. DFA consists of five steps:

- a random walk process is constructed from the time series:

$$u(n) = \sum_{k=1}^{n} (x_k - \overline{x}), \;\; n = 1, 2, \cdots, N, \tag{1}$$

- divide the random walk process into non-overlapping segments
- determine the local trends of each segment as the best polynomial fit
- extract variance of the differences between the random walk process and the local trends
- determine average variance over all the segment

Adaptive fractal analysis

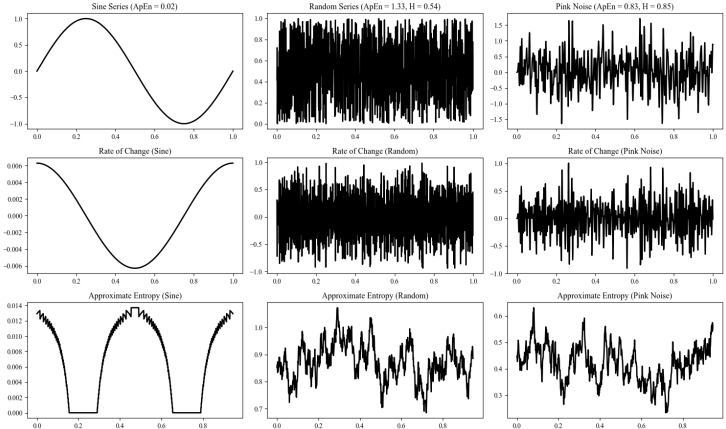
DFA suffers catastrophically from discontinuities at the boundaries of adjacent segments, non-stationarity, and nonlinear oscillatory components. AFA identifies a global smooth trend by optimally combining local linear or polynomial fitting – nonlinear adaptive multi-scale decomposition algorithm

- partition into overlapping segments of length $w = 2n + 1$, where neighboring segments overlap by $n + 1$ points.
- fit best polynomial of order M (OLS), combine polynomials in overlapping regions to a single global smooth trend ($y^{i}(l_1)$ and $y^{(i+1)}(l_2)$, where $l_1, l_2 = 1, \cdots, 2n + 1$,).

$$y^{(c)}(l) = w_1 y^{(i)}(l + n) + w_2 y^{(i+1)}(l), \;\; l = 1, 2, \cdots, n + 1, \tag{2}$$

- $w_1 = \left(1 - \frac{l-1}{n}\right)$ and $w_2 = \frac{l-1}{n}$ can be written as $(1 - d_j/n)$ for $j = 1, 2$, and d_j is the distances between the point and the centers of $y^{(i)}$ and $y^{(i+1)}$
- For window size w and random walk $u(i)$, a global trend $v(i)$, $i = 1, 2, \cdots, N$, where N is the length of the walk, the residual of the fit, $u(i) - v(i)$, characterizes fluctuations around the global trend, and its variance yields the Hurst parameter H according to the following scaling equation:

$$F(w) = \Big[\frac{1}{N}\sum_{i=1}^{N}(u(i) - v(i))^2\Big]^{1/2} \sim w^H. \tag{3}$$


Sine Series (ApEn = 0.02)
Random Series (ApEn = 1.33, H = 0.54)
Pink Noise (ApEn = 0.83, H = 0.85)
Rate of Change (Sine)
Rate of Change (Random)
Rate of Change (Pink Noise)
Approximate Entropy (Sine)
Approximate Entropy (Random)
Approximate Entropy (Pink Noise)

Approximate entropy

For series $X = x(1), \ldots, x(n)$, sub-sequences of length m, $y_i^m = [x(i), \ldots, x(i + (m - 1))]$, and tolerance r, *Approximate Entropy* ($ApEn$) is estimated by computing the Chebyshev distance between each sub-sequence y_i^m and y_j^m

$$d_{i,j}^m = \max_k \mid y_i^m - y_j^m(k) \mid$$

for each sub-sequence y_i^m to compute the count C

$$C_i^m(r) = \frac{1}{n - m + 1} \sum_{j=1}^{n-m+1} H(r - d_{i,j}^m)$$

where $H(.)$ is the Heaviside function

$$H(x) = \begin{cases} 1, & x > 0 \\ 0, & x \leq 0 \end{cases}$$

then define

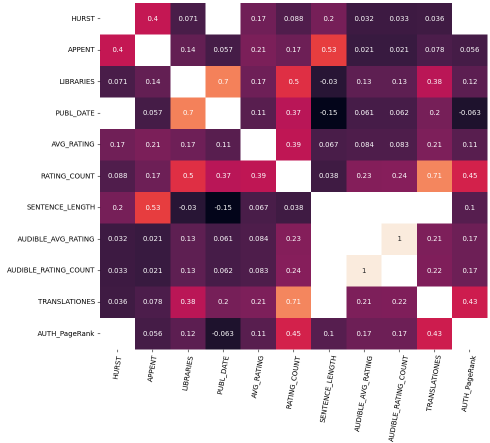
$$\phi^m(r) = \frac{1}{n - m + 1} \sum_{i=1}^{n-m+1} \log(C_i^m(r))$$

where $\log(.)$ is the natural logarithm. Repeat the above for all sub-sequences of length $m + 1$ to compute $\phi^(m + 1)(r)$, then Approximate Entropy is

$$ApEn(m, r) = \phi^m(r) - \phi^{m+1}(r)$$

OUTLINE

1. Introduction
2. Methods
3. **Results**
4. Discussion


HURST
APPENT
LIBRARIES
PUBL_DATE
AVG_RATING
RATING_COUNT
SENTENCE_LENGTH
AUDIBLE_AVG_RATING
AUDIBLE_RATING_COUNT
TRANSLATIONES
AUTH_PageRank
1.0
0.8
0.6
0.4
0.2
0.0
-0.2

OUTLINE

1. Introduction
2. Methods
3. Results
4. Discussion

In summary

- The narrative structure of successful literature displays a tension between local uncertainty (i.e., segments of successful arcs show higher approximate entropy) and global coherence (i.e., successful arcs display long-range dependencies).
- The tension between the local and global organization of language reflects that literature, like other cultural artifacts, has to balance attention and motivation to be culturally transmitted successfully [1, 2].

THANK YOU

kln@cas.au.dk
chc.au.dk

SLIDES

knielbo.github.io/files/kln-cudan23-1.pdf

[1] Hu, Q., Liu, B., Thomsen, M. R., Gao, J., and Nielbo, K. L. (2020). Dynamic evolution of sentiments in never let me go: Insights from multifractal theory and its implications for literary analysis. *Digital Scholarship in the Humanities*, 36(2):322–332.

[2] Mohseni, M., Gast, V., and Redies, C. (2021). Fractality and variability in canonical and non-canonical english fiction and in non-fictional texts. *Frontiers in Psychology*, 12.