On the Distillation of Stories for Transferring Narrative Arcs in Collections of Independent Media

Dylan R. Ashley [1, 2, 3 *]
Vincent Herrmann [1, 2, 3 *]
Zachary Friggstad [4]
Jürgen Schmidhuber [1, 2, 3, 5, 6]

[1] Dalle Molle Institute for Artificial Intelligence Research, Lugano, Switzerland
[2] Università della Svizzera italiana, Lugano, Switzerland
[3] Scuola universitaria professionale della Svizzera italiana, Lugano, Switzerland
[4] University of Alberta, Edmonton, Canada
[5] NNAISENSE, Lugano, Switzerland
[6] AI Initiative, King Abdullah University of Science and Technology, Thuwal, Saudi Arabia
[*] *Equal Contribution*

Abstract

The act of telling stories is a fundamental part of what it means to be human. This work introduces the concept of narrative information, which we define to be the overlap in information space between a story and the items that compose the story. Using contrastive learning methods, we show how modern artificial neural networks can be leveraged to distill stories and extract a representation of the narrative information. We then demonstrate how evolutionary algorithms can leverage this to extract a set of narrative templates and how these templates–in tandem with a novel curve-fitting algorithm we introduce–can reorder music albums to automatically induce stories in them. In the process of doing so, we give strong statistical evidence that these narrative information templates are present in existing albums. While we experiment only with music albums here, the premises of our work extend to any form of (largely) independent media.

- Decomposition of a story into atoms and narrative structure
- Atoms can be words, images, music tracks, etc.
- Narrative structure is any meaningful arrangement of the atoms
- A collection of atoms, together with a meaningful arrangement (narrative structure) is a story
- Narrative information of an atom is its mutual information with the narrative structure

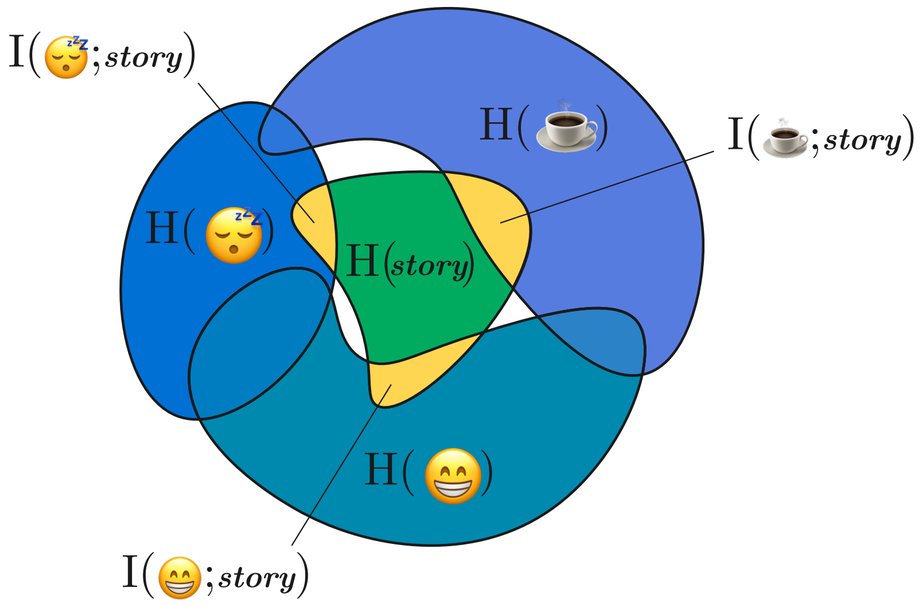
$$X = \{\mathbf{x}_1, \quad \mathbf{x}_2, \quad \mathbf{x}_3\}$$

$$\{\text{😴}, \text{☕}, \text{😀}\}$$ Atoms

Structure

story

$$\mathbf{s} = \{(\mathbf{x}_1, 1), (\mathbf{x}_2, 2), (\mathbf{x}_3, 3)\}$$

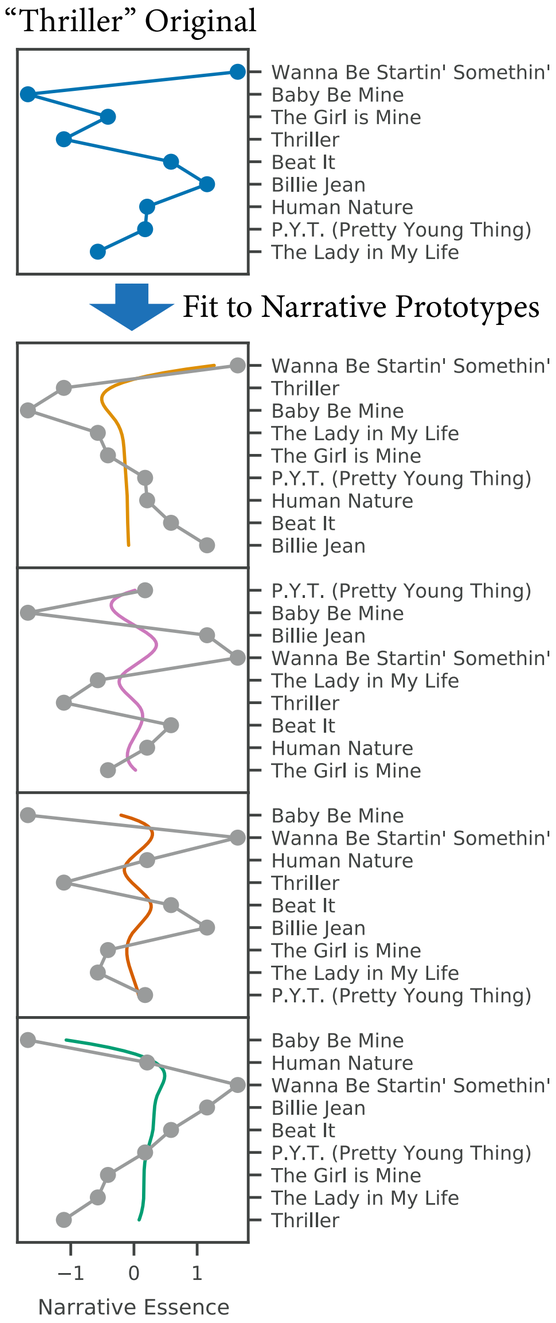


$$f_E := \arg\max_f \mathrm{I}\left(\{f(\mathbf{x}) | \mathbf{x} \in X\}; \mathbf{s}\right)$$

$$\mathbf{z}_k := f_E\left(\mathbf{x}_k\right)$$

- Narrative essence is a concrete low-dimensional representation of the latent property of the atoms that is most informative about the narrative structure
- A neural network-based feature extractor can be learned from data using noise contrastive estimation (InfoNCE)

- Using a genetic algorithm, we can discover narrative prototypes in the data
- A prototypical story can be induced into a music album by ordering it such that its narrative essence shape matches one of the prototypical shapes



arXiv

GitHub

Feature extractor f_θ

Sequence model g_ϕ

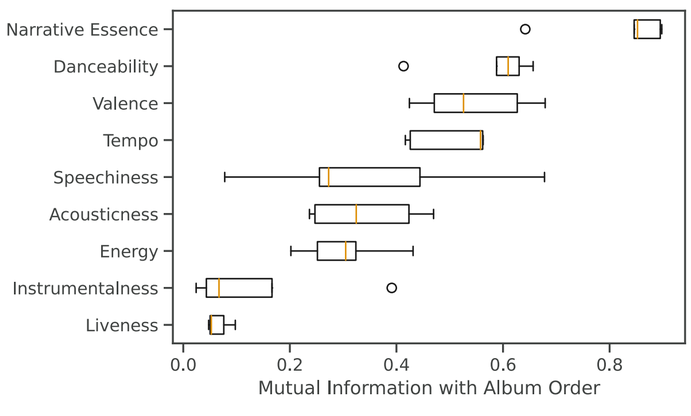
Correct sequence
$s^* = (f_\theta(\mathbf{x}_1), f_\theta(\mathbf{x}_2), f_\theta(\mathbf{x}_3), ...)$

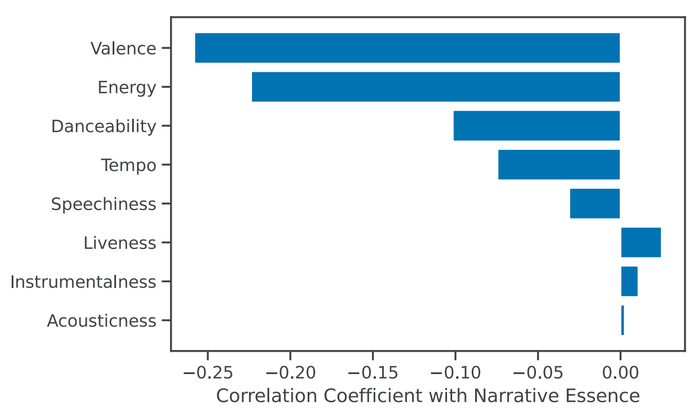
Set of permutations of s^*: S

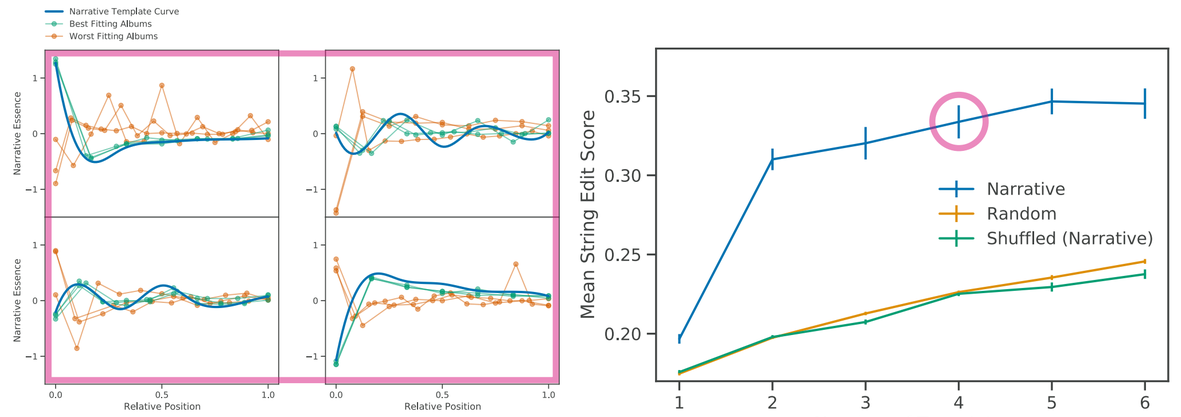
InfoNCE loss
$\mathcal{L}_{\mathrm{N}}(\theta, \phi; \mathcal{D}) = -\mathbb{E}_{S \sim \mathcal{D}}\left[\log \frac{g_\phi(s^*)}{\sum_{s \in S} g_\phi(s)}\right]$

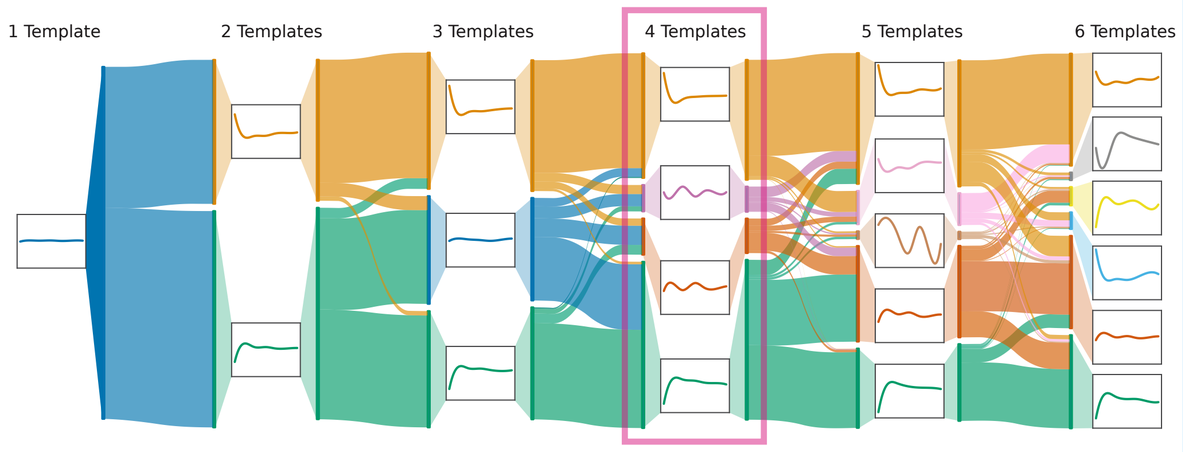
Table 1: Mutual Information (in bits) on the FMA validation set for different dimensionalities of narrative essence. Results are from five runs.

Features	Mutual Information
1	1.924 ± 0.0296
2	1.936 ± 0.0183
4	1.957 ± 0.0217
8	1.950 ± 0.0216
16	1.975 ± 0.0150









Istituto Dalle Molle di studi sull'intelligenza artificiale
IDSIA
Università della Svizzera italiana
Scuola universitaria professionale della Svizzera italiana
SUPSI
UNIVERSITY OF ALBERTA
KAUST
European Research Council
Established by the European Commission
CSCS
Centro Svizzero di Calcolo Scientifico
Swiss National Supercomputing Centre