

January 15, 2025

# Mathematical Framework for Narrative Synthesis using Latent Space Techniques

## 1. Formal Definition of Narratives in Latent Space:

A narrative  $\mathcal{N}$  can be represented in a latent space  $\mathcal{L} \subseteq \mathbb{R}^d$ , where  $d$  is the dimensionality of the latent space. We consider two primary representations:

- *Narrative as a Point*: A static representation where the narrative is summarized by a single point in the latent space:

$$\mathcal{N} = \mathbf{n} \in \mathcal{L}$$

where  $\mathbf{n}$  is a  $d$ -dimensional vector.

- *Narrative as a Trajectory*: A dynamic representation where the narrative unfolds as a path or trajectory in the latent space, parameterized by a time or thematic progression variable  $t \in [0, T]$ :

$$\mathcal{N} = \gamma(t) : [0, T] \rightarrow \mathcal{L}$$

where  $\gamma(t)$  is a vector-valued function in  $\mathbb{R}^d$ .

## 2. Mathematical Formulation for Chirality between Narratives:

Chirality refers to the handedness or mirror asymmetry between narratives. We can define it differently for point and trajectory representations:

- *Chirality between two narratives as points* ( $\mathbf{n}_1, \mathbf{n}_2$ ): Consider a reference vector  $\mathbf{r}$  that defines a plane or direction of comparison. Chirality can be assessed based on the orientation of the triplet  $(\mathbf{r}, \mathbf{n}_1, \mathbf{n}_2)$ . For example, in 3D:

$$\text{Chirality}(\mathbf{n}_1, \mathbf{n}_2 | \mathbf{r}) = \text{sign}((\mathbf{n}_1 - \mathbf{r}) \times (\mathbf{n}_2 - \mathbf{r}) \cdot \mathbf{r})$$

where  $\times$  is the cross product and  $\cdot$  is the dot product. A non-zero value indicates chirality. Generalization to higher dimensions might involve the sign of a higher-order determinant.

- *Chirality between two narratives as trajectories* ( $\gamma_1(t), \gamma_2(t)$ ): This can involve comparing the orientation of their tangent vectors at corresponding points. Let  $\dot{\gamma}_1(t) = \frac{d\gamma_1(t)}{dt}$  and  $\dot{\gamma}_2(t) = \frac{d\gamma_2(t)}{dt}$ . Considering a reference trajectory tangent  $\dot{\mathbf{r}}(t)$ :

$$\text{Chirality}(\gamma_1, \gamma_2 | \dot{\mathbf{r}}(t)) = \int_0^T \text{sign}((\dot{\gamma}_1(t) - \dot{\mathbf{r}}(t)) \times (\dot{\gamma}_2(t) - \dot{\mathbf{r}}(t)) \cdot \dot{\mathbf{r}}(t)) dt$$

Integrating over the trajectory provides a global measure of chirality.

## 3. Mathematical Formulation for Orthogonality between Narratives:

Orthogonality signifies a high degree of dissimilarity or independence between narratives.

- *Orthogonality between two narratives as points* ( $\mathbf{n}_1, \mathbf{n}_2$ ):

$$\text{Orthogonality}(\mathbf{n}_1, \mathbf{n}_2) = \frac{|\mathbf{n}_1 \cdot \mathbf{n}_2|}{\|\mathbf{n}_1\| \|\mathbf{n}_2\|}$$

A value close to 0 indicates high orthogonality.

- *Orthogonality between two narratives as trajectories* ( $\gamma_1(t), \gamma_2(t)$ ): We can consider the orthogonality of their average vectors or the orthogonality of their tangent vectors over time:

$$\text{Orthogonality}(\gamma_1, \gamma_2) = \frac{\left| \int_0^T \gamma_1(t) dt \cdot \int_0^T \gamma_2(t) dt \right|}{\left\| \int_0^T \gamma_1(t) dt \right\| \left\| \int_0^T \gamma_2(t) dt \right\|}$$

or, considering tangent vectors:

$$\text{Orthogonality}_{\text{tangent}}(\gamma_1, \gamma_2) = \frac{1}{T} \int_0^T \frac{|\dot{\gamma}_1(t) \cdot \dot{\gamma}_2(t)|}{\|\dot{\gamma}_1(t)\| \|\dot{\gamma}_2(t)\|} dt$$

#### 4. Algorithms or Formulas for Narrative Synthesis:

Given a set of narratives in the latent space, we can synthesize new narratives using various operations:

- *Linear Combination (for point narratives)*: Given narratives  $\mathbf{n}_1, \mathbf{n}_2, \dots, \mathbf{n}_k$  and weights  $w_1, w_2, \dots, w_k$  where  $\sum_{i=1}^k w_i = 1$ :

$$\mathbf{n}_{\text{synth}} = \sum_{i=1}^k w_i \mathbf{n}_i$$

- *Interpolation (for trajectory narratives)*: Given trajectories  $\gamma_1(t)$  and  $\gamma_2(t)$  and a blending function  $\alpha(t) \in [0, 1]$ :

$$\gamma_{\text{synth}}(t) = (1 - \alpha(t))\gamma_1(t) + \alpha(t)\gamma_2(t)$$

- *Extrapolation (for point narratives)*: Given two narratives  $\mathbf{n}_1$  and  $\mathbf{n}_2$ , and an extrapolation factor  $\lambda > 1$ :

$$\mathbf{n}_{\text{synth}} = \mathbf{n}_1 + \lambda(\mathbf{n}_2 - \mathbf{n}_1)$$

- *Vector Addition (for point narratives)*:

$$\mathbf{n}_{\text{synth}} = \mathbf{n}_1 + \mathbf{n}_2$$

- *Concatenation (for trajectory narratives)*: Given two trajectories  $\gamma_1(t)$  and  $\gamma_2(t)$ , the synthesized trajectory can be:

$$\gamma_{\text{synth}}(t) = \begin{cases} \gamma_1(2t) & 0 \leq t \leq 0.5 \\ \gamma_2(2t - 1) & 0.5 < t \leq 1 \end{cases}$$

(with potential smoothing or transitions).

- *Synthesis Guided by Chirality and Orthogonality*: We can synthesize narratives that maximize or minimize specific properties. For example, find a narrative  $\mathbf{n}_{\text{synth}}$  such that its average orthogonality to a set of existing narratives is below a threshold, and its chirality with respect to a target narrative has a desired sign. This can be framed as an optimization problem.

#### 5. Additional Relevant Equations and Concepts:

- *Distance Metric in Latent Space:* The choice of distance metric is crucial. Common metrics include Euclidean distance, cosine similarity, and Manhattan distance.

$$d(\mathcal{N}_1, \mathcal{N}_2) = \|\mathbf{n}_1 - \mathbf{n}_2\|_p \quad (\text{for point narratives, where } p \text{ is the norm})$$

For trajectory narratives, Dynamic Time Warping (DTW) or Fréchet distance can be used.

- *Latent Space Transformation:* Applying transformations to the latent space can affect narrative properties. For example, rotation matrices can alter chirality.

$$\mathbf{n}' = R\mathbf{n}$$

where  $R$  is a rotation matrix.

- *Narrative Density Estimation:* Understanding the distribution of narratives in the latent space using techniques like Kernel Density Estimation (KDE).

$$\hat{f}(\mathbf{x}) = \frac{1}{nh^d} \sum_{i=1}^n K\left(\frac{\mathbf{x} - \mathbf{n}_i}{h}\right)$$

where  $K$  is a kernel function,  $h$  is the bandwidth.

- *Clustering of Narratives:* Grouping similar narratives using algorithms like k-means or DBSCAN.
- *Dimensionality Reduction Techniques:* If the initial latent space is high-dimensional, techniques like Principal Component Analysis (PCA) or t-SNE can be used to simplify the space while preserving important narrative features.

$$\mathbf{n}_{\text{reduced}} = W^T \mathbf{n}$$

where  $W$  is the matrix of principal components.

This framework provides a mathematical foundation for understanding and synthesizing narratives within a latent space, leveraging concepts like chirality and orthogonality to control the properties of the generated narratives. The specific implementation and choice of formulas will depend on the nature of the narratives and the desired synthesis outcomes.