Decentralized Autonomous Narrative Networks (DANN)
A Multi-Model Framework for Reinforcement Learning with Veracity, Influence, and Reputation

Paul Lowndes
ZeroTrust@NSHkr.com

January 1, 2025

Abstract

This paper introduces Decentralized Autonomous Narrative Networks (DANN), a novel multi-model framework for reinforcement learning inspired by the complexities of human social interactions, information asymmetry, and the manipulation of narratives. Unlike traditional multi-agent reinforcement learning (MARL) systems, DANN equips each agent with an independent, internal model, analogous to a Large Concept Model (LCM), which maintains the agent's unique knowledge, beliefs, and evolving narrative represented as a sequence of concept embeddings. We propose a decentralized approach to veracity, where "ground truth" emerges from the interactions and consensus mechanisms within the network, rather than being imposed by a central authority. The framework incorporates agent-specific parameters, narrative-based reward functions, and mechanisms for cross-narrative influence and agent switching to model the dynamic and often adversarial nature of real-world information ecosystems. We further explore the use of concepts from topology, game theory, and information theory to mathematically formalize the notions of narrative divergence, asymmetry thresholds, and the influence of agents on one another. Through this work, we aim to develop a more robust and ethically aware approach to AI development, capable of modeling the complexities of deception, manipulation, and the social construction of reality, while also offering potential insights into the mitigation of these phenomena. The DANN framework, while still theoretical, provides a foundation for future research into the development of AI systems that can navigate and potentially counteract the weaponization of information and ultimately, foster a more transparent and equitable information environment. This also, of course, has implications for national security, individual privacy, and the potential development of novel and effective methods of online harassment, all of which must be taken into consideration. This framework also potentially provides a new tool for understanding, and even countering, the spread of misinformation. It also opens up entirely new avenues for potential abuse.

1 Introduction

The proliferation of artificial intelligence systems and their increasing role in shaping information flows has created new challenges in understanding and managing narrative dynamics in digital spaces. Traditional multi-agent reinforcement learning (MARL) approaches often fail to capture the nuanced interplay between agents' beliefs, knowledge, and the narratives they construct and propagate. This paper introduces Decentralized Autonomous Narrative Networks (DANN), a framework that explicitly models these dynamics through a combination of embedding spaces, belief systems, and narrative evolution mechanisms.

2 Framework Overview

2.1 Core Components

We begin by defining the fundamental mathematical structures that underpin the DANN framework:

Definition 1 (Embedding Space)**.** *The global embedding space E_G is a metric space (E_G, d) where:*

- $d : E_G \times E_G \to \mathbb{R}_{\geq 0}$ *is a distance function*
- *For all* $x, y \in E_G$*:* $d(x, y) = 0 \iff x = y$ *(identity)*
- *For all* $x, y \in E_G$*:* $d(x, y) = d(y, x)$ *(symmetry)*
- *For all* $x, y, z \in E_G$*:* $d(x, z) \leq d(x, y) + d(y, z)$ *(triangle inequality)*

Definition 2 (Agent Space)**.** *For each agent a_i, its local embedding space $E_i \subseteq E_G$ is equipped with:*

- *Knowledge set* $K_{i,t} \subset E_i$ *at time* t
- *Belief set* $B_{i,t} \subset E_i$ *at time* t
- *Narrative sequence* $N_{i,t} = (c_{i,1}, c_{i,2}, \ldots, c_{i,T}) \in E_i^T$

where $K_{i,t} \subseteq B_{i,t}$ (knowledge is a subset of beliefs).

3 Mathematical Framework

3.1 Veracity Function Properties

The veracity function $V : E_G \to [0, 1]$ satisfies:

Property 1 (Veracity Axioms)**.** *For all $x, y \in E_G$:*

- $V(x) = 1 \iff x \in T$ *(truth region)*
- $\|x - y\| \leq \epsilon \implies |V(x) - V(y)| \leq \delta$ *(continuity)*
- $V(x) = 0 \implies x$ *is maximally inconsistent with truth*

3.2 Narrative Dynamics

Definition 3 (Narrative Divergence)**.** *The divergence D between narratives satisfies:*

$$D(N_{i,t}, N_{j,t}) = \sum_{k=1}^{T} w(c_{i,k}) \cdot d(c_{i,k}, c_{j,k}) \tag{1}$$

where $w(c) = f(V(c))$ for some monotonic function $f : [0, 1] \to [0, 1]$.

3.3 Agent Interaction Mechanisms

3.3.1 Knowledge Propagation

Knowledge updates follow:

$$K_{i,t+1} = K_{i,t} \cup \{e \in E_i \mid V(e,T) > \tau_K \wedge \exists j : e \in K_{j,t}\} \tag{2}$$

where τ_K is the knowledge acceptance threshold.

3.3.2 Belief Evolution

Belief updates incorporate both knowledge and social influence:

$$B_{i,t+1} = f_B(B_{i,t}, K_{i,t+1}, \sum_{j \neq i} \alpha_{ij} B_{j,t}) \tag{3}$$

where α_{ij} represents the influence weight of agent j on agent i.

4 Learning Mechanisms

4.1 Narrative-Based Reward

The reward function combines environmental and narrative quality:

$$R_i(s_t, a_t, s_{t+1}) = \alpha \cdot R_{\text{env}}(s_t, a_t, s_{t+1}) + \beta \cdot Q(N_{i,t+1}) \tag{4}$$

where:

- $Q(N) = \gamma_1 C(N) + \gamma_2 V_{\text{avg}}(N) + \gamma_3 I(N)$
- $C(N)$ measures narrative coherence
- $V_{\text{avg}}(N)$ is the average veracity
- $I(N)$ measures narrative influence

4.2 Agent-Switching Mechanism

The switching function is defined as:

$$S_i(t) = \arg\max_j \{Q(M_{i,j}, N_{i,t}, \text{Context}_t) + \lambda H(j)\} \tag{5}$$

where:

- $H(j)$ is an entropy term promoting exploration
- λ balances exploitation vs. exploration
- Context_t includes environmental and social factors

5 Alternative Methods

5.1 Veracity Function

$$V(e, T, a_i, C) = w_1 \cdot \left(1 - \frac{d(e, T)}{\max_{x \in E_G} d(x, T)}\right) + w_2 \cdot S_R(\text{Source}(e)) + w_3 \cdot C_A(e, C) + w_4 \cdot D_R(e, a_i) \quad (6)$$

where:
$S_R(e) = \alpha \cdot H(\text{Source}(e)) + \beta \cdot E(\text{Source}(e)) + \gamma \cdot (1 - B(\text{Source}(e))) + \delta \cdot \text{Corroboration}(e)$
$H(s) =$ historical accuracy of source s
$E(s) =$ expertise level of source s
$B(s) =$ detected biases of source s
$C_A(e, C) =$ consistency and coherence of e within context C
$D_R(e, a_i) =$ assessed defamation risk of e towards agent a_i
$\text{Corroboration}(e) =$ measure of agreement with independent sources (7)

5.2 Narrative Divergence

$$D(N_{i,t}, N_{j,t}) = \sum_{k=1}^{T} w(c_{i,k}) \cdot d(c_{i,k}, c_{j,k}) \quad (8)$$

where:
$$w(c) = \frac{1}{1 + e^{-V(c,T,a_i,C)}} \quad \text{(sigmoid function applied to veracity)} \quad (9)$$

This uses a sigmoid function to transform the veracity score into a weight, emphasizing embeddings with higher veracity.

5.3 Influence Weighting

$$\alpha_{ij}(t) = \sigma(\beta_1 \cdot N_{ij} + \beta_2 \cdot \text{Rep}_j(t) + \beta_3 \cdot E_j + \beta_4 \cdot P_j) \quad (10)$$

where:
$N_{ij} =$ strength of network connection between a_i and a_j
$\text{Rep}_j(t) =$ reputation score of agent a_j at time t
$E_j =$ domain expertise of agent a_j
$P_j =$ platform-specific influence metrics for agent a_j
$\sigma =$ sigmoid function
$\beta_1, \beta_2, \beta_3, \beta_4$ are learned parameters (11)

This formula models the influence weight as a function of network connections, historical reliability (reputation), expertise, and platform-specific factors, all combined through a sigmoid function to produce a value between 0 and 1.

5.4 Narrative-Based Reward Function

$$R_i(s_t, a_t, s_{t+1}) = \alpha \cdot R_{\text{env}}(s_t, a_t, s_{t+1}) + \beta \cdot Q(N_{i,t+1}) \tag{12}$$

where:

$$Q(N_{i,t}) = \gamma_1 \cdot C(N_{i,t}) + \gamma_2 \cdot \frac{1}{|N_{i,t}|} \sum_{c \in N_{i,t}} V(c, T, a_i, C_t) + \gamma_3 \cdot \sum_{j \neq i} \alpha_{ji}(t) \cdot (1 - D(N_{i,t}, N_{j,t})) \tag{13}$$

$C(N_{i,t})$: Coherence of the narrative $N_{i,t}$ (to be defined further, potentially based on average distance between embeddings or logical consistency). $V_{avg}(N_{i,t})$: Average veracity of concept embeddings in the narrative. $I(N_{i,t})$: Influence of the narrative, here modeled as a function of how much it reduces divergence with other agents' narratives, weighted by their influence. This could be further refined to include, for example, how much closer an agent's narrative moves toward the "ground truth" after receiving input from another agent.

5.5 Agent Switching

$$S_i(t) = \underset{j \in \{1,\ldots,k\}}{\arg\max} \{Q(M_{i,j}, N_{i,t}, \text{Context}_t) + \lambda \cdot H(j)\} \tag{14}$$

where:

$$\begin{aligned} &H(j) = -\sum_{t'=1}^{t-1} p(j|t') \log p(j|t') \quad \text{(entropy of model selection history)} \\ &p(j|t') = \text{probability of selecting model } j \text{ at time } t' \\ &\lambda = \text{exploration coefficient} \end{aligned} \tag{15}$$

This formulation of $H(j)$ encourages exploration by favoring models that haven't been selected as often in the past, based on the selection history up to the current time step. This would encourage the system to explore its options more thoroughly, even potentially selecting a model that does not maximize performance but is useful for other reasons, such as for gathering more information, if applicable.

6 Discussion and Future Work

[This section would discuss implications, limitations, and future research directions]