# Shadow Cones: A Generalized Framework for Partial Order Embeddings


Tao Yu tyu@cs.cornell.edu
Toni J.B. Liu jl3499@cornell.edu
Albert Tseng albert@cs.cornell.edu
Christopher De Sa cdesa@cs.cornell.edu



Hyperbolic space are well-suited for embedding hierarchies, such as trees. We propose to embed **partial orders** as subset relations between shadows formed by a light source and opaque objects in Riemannian space. This framework, termed **Shadow Cones**, enables use to generalize hyperbolic entailment cones. Our experiments on datasets of various sizes and hierarchical structures show that shadow cones consistently and significantly outperform the existing entailment cone constructions.


## Primer on hyperbolic space

Hierarchical and graphical data, such as biological phylogenetic trees and social networks, are more naturally modeled in hyperbolic space. This is because in hyperbolic space, the volume of a ball grows exponentially for large radius, which matches the number of nodes in a tree; in contrast, this volume grows only polynomially in Euclidean space. This work uses two models of hyperbolic space: Poincaré ball and half-space.

The **Poincaré ball** is given by $\mathcal{B}^n = \{\boldsymbol{x} \in \mathbb{R}^n : \|\boldsymbol{x}\| < 1/\sqrt{k}\}$

Distances on $\mathcal{B}^n$ are defined as

$$d_{\mathcal{B}}(\boldsymbol{x}, \boldsymbol{y}) = \frac{1}{\sqrt{k}} \operatorname{arcosh}\left(1 + 2\frac{k\|\boldsymbol{x}-\boldsymbol{y}\|^2}{(1-k\|\boldsymbol{x}\|^2)(1-k\|\boldsymbol{y}\|^2)}\right)$$

The **Poincaré half-space** is given by $\mathcal{U}^n = \{\boldsymbol{x} \in \mathbb{R}^n : x_n > 0\}$.

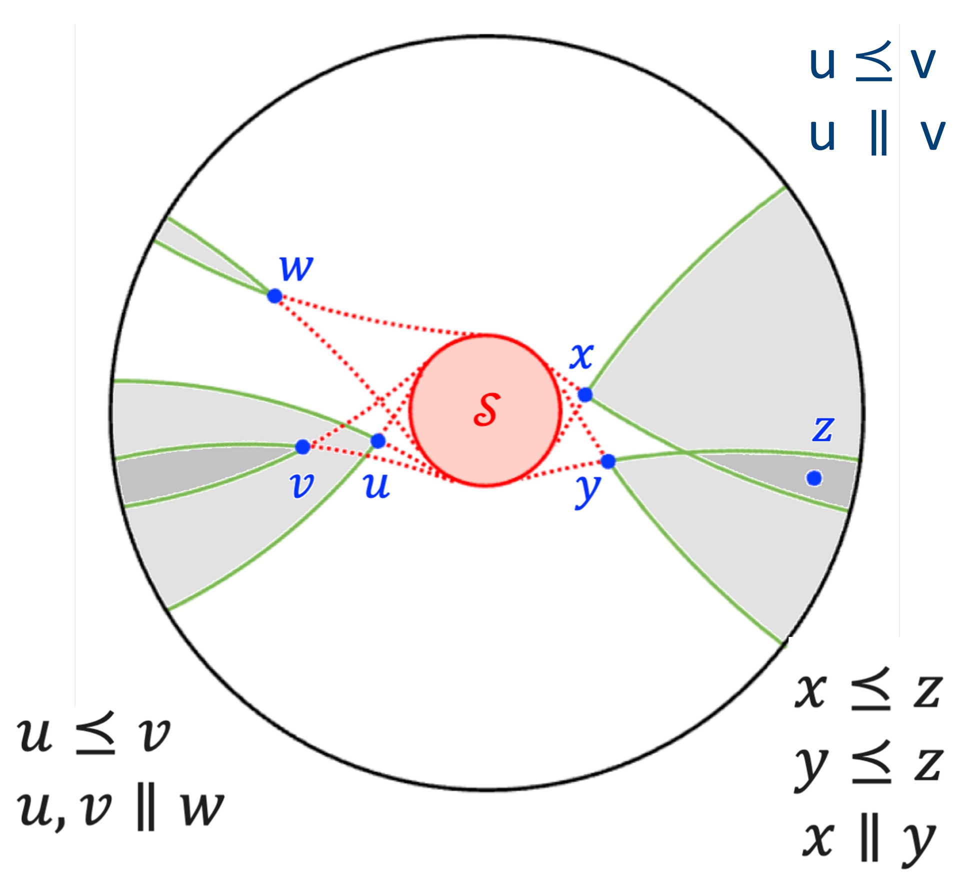
Distances on $\mathcal{U}^n$ are defined as

$$d_{\mathcal{U}}(\boldsymbol{x}, \boldsymbol{y}) = \frac{1}{\sqrt{k}} \operatorname{arcosh}\left(1 + \frac{\|\boldsymbol{x}-\boldsymbol{y}\|^2}{2x_n y_n}\right)$$

## Representing partial order as shadow casting

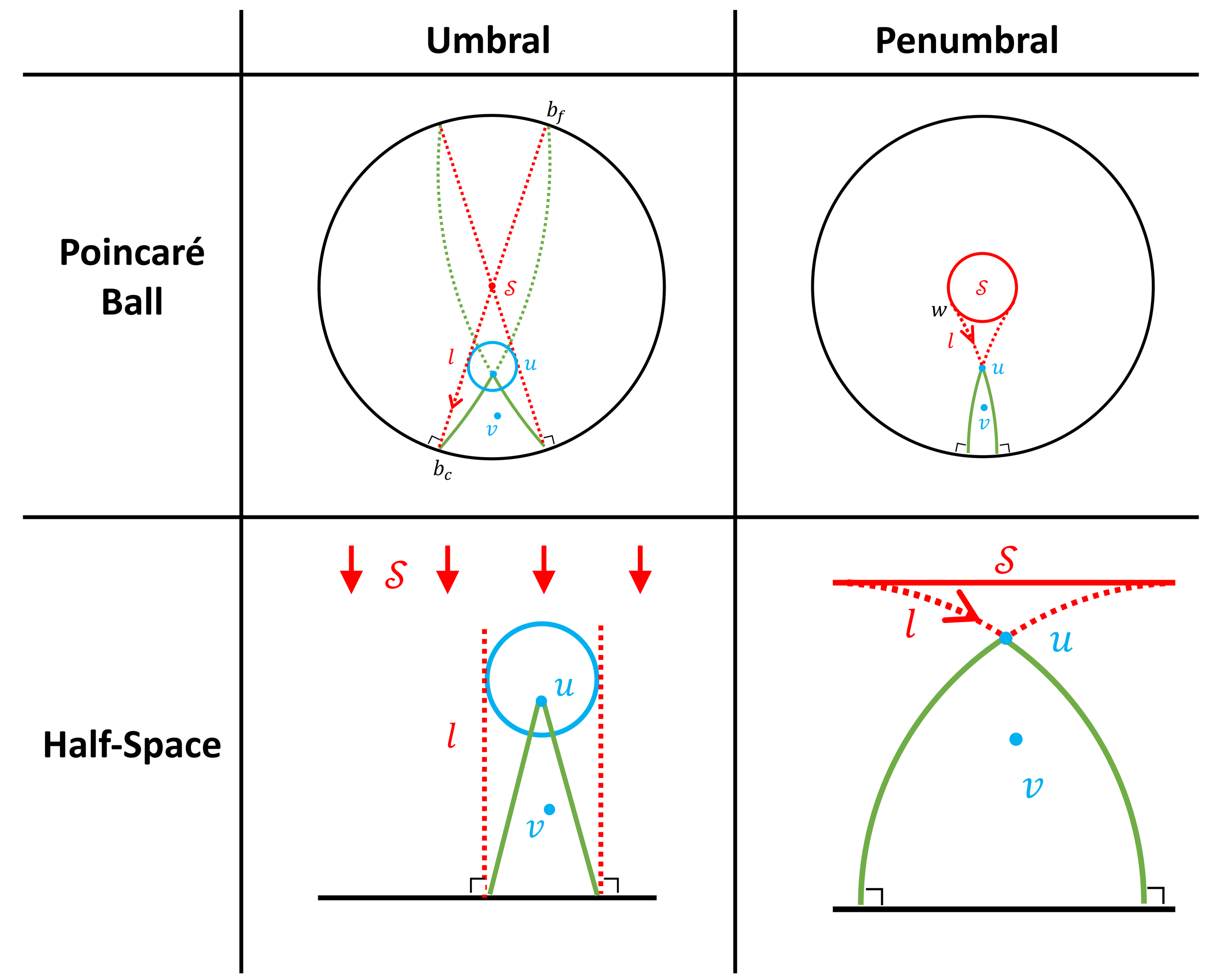
The shadow cones framework geometrically represents the partial relation using subset relation between shadows. v is represented as a child of u, when the shadow cast by v is a subset of the shadow cast by u.

$u \preceq v$ v is a child of u
$u \parallel v$ v is incomparable to u



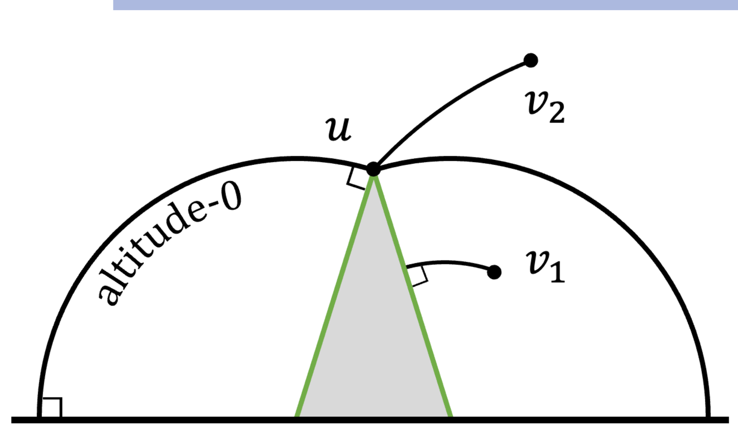

Link to full paper

## 4 formulations of hyperbolic shadow cones



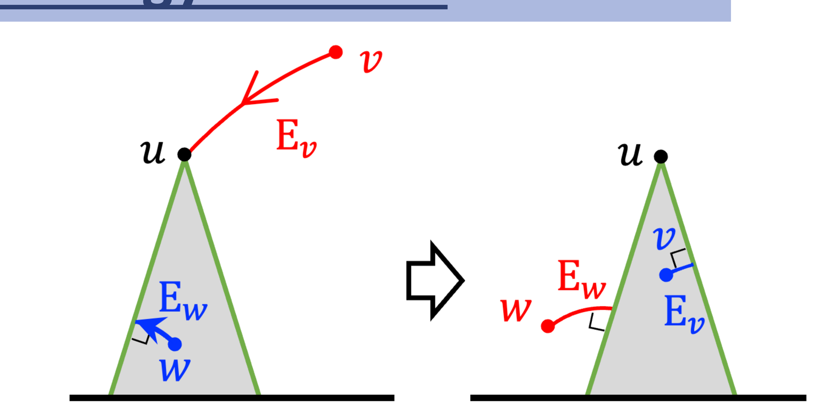

| Cone | Model | Form. # | Emb. Type | Convex? | Light Source | $\theta_u$ |
|---|---|---|---|---|---|---|
| Umbral | Half-Space | 1 | Ball | No | Point at $\infty$ | Fixed |
| | Poincaré Ball | 2 | Ball | No | Point at Origin | Varying |
| Penumbral | Poincaré Ball | 3 | Point | Yes | Ball at Origin | Varying |
| | Half-Space | 4 | Point | Yes | Horosphere | Varying |

## Distance-based energy function




Shortest distances to u's cone, as calculated from points above and below the 0-altitude line.

*The penumbral shadow cones in the Poincaré ball is identical to Ganea's entailment cones (Ganea et al. 2018)*




Toy example: $u \preceq v$, $u \parallel w$

Left: v is a child of u, but wrongly initialized outside of u's cone. w is incomparable with u, but initialized inside the cone.

Right: The energy gradients pulls v inside the cone, and pushes w outside.

Red: positive; Blue: negative

## Contrastive training loss

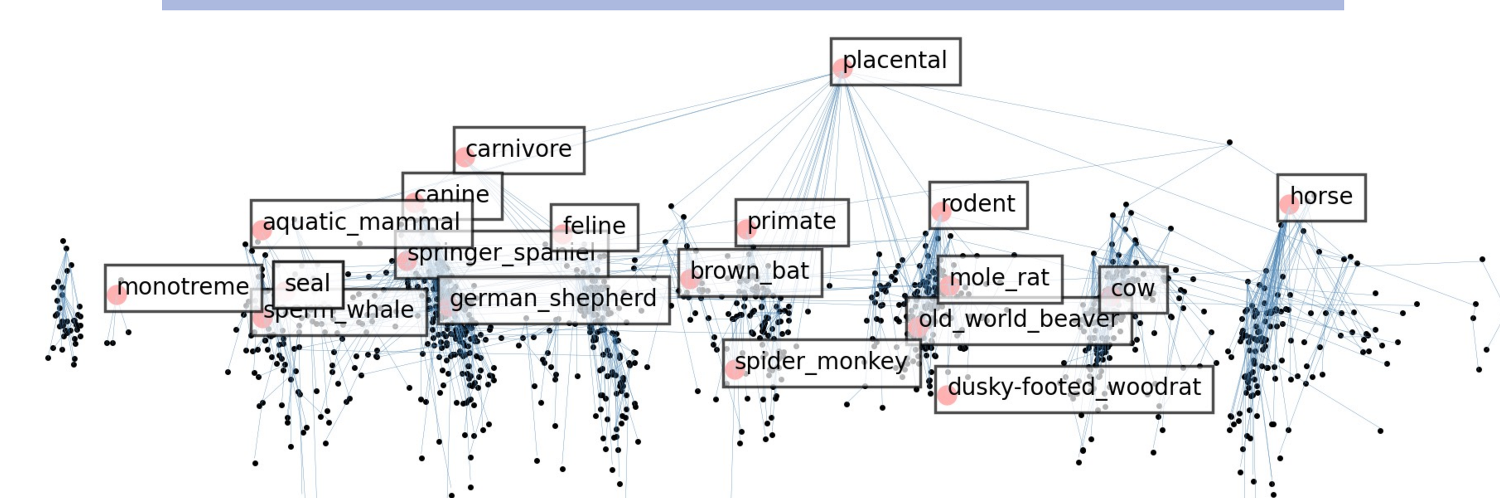
If $u \preceq v$, pull them together until their cones are correctly nested.
If $u \parallel v$, push them apart until their cones are not nested.

$$\mathcal{L}_{\gamma_1,\gamma_2} = \sum_{(\boldsymbol{u},\boldsymbol{v})\in P} \log \frac{\exp(-\max(E(\boldsymbol{u},\boldsymbol{v}),\gamma_2))}{\sum_{(\boldsymbol{u}',\boldsymbol{v}')\in N} \exp(\max(\gamma_1 - E(\boldsymbol{u}',\boldsymbol{v}'),0))}$$

P: the edge set of positive relations
N: that of negative relations
E(u, v) = d(v, Cone(u)) is the two-case distance

This loss allows us to choose how far to push negative samples away from the cone (distance $\gamma_1 > 0$), and how deep to pull positive samples into the cone (distance $\gamma_2 > 0$).

## Results and benchmarks




Mammal embedding using umbral shadow cones in 2D half-space

F1 score (%) on mammal sub-graph with best numbers **bolded**

| Non-basic-edge Percentage | Dimension = 2 | | | | | Dimension = 5 | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | 0% | 10% | 25% | 50% | 90% | 0% | 10% | 25% | 50% | 90% |
| GBC-box | 23.4 | 25.0 | 23.7 | 43.1 | 48.2 | 35.8 | 60.1 | 66.8 | 83.8 | **97.6** |
| VBC-box | 20.1 | 26.1 | 31.0 | 33.3 | 34.7 | 30.9 | 43.1 | 58.6 | 74.9 | 69.3 |
| Entailment Cone | 54.4 | 61.0 | 71.0 | 66.5 | 73.1 | 56.3 | 81.0 | **84.1** | 83.6 | 82.9 |
| Umbral-half-space | **57.7** | 73.7 | **77.4** | **80.3** | **79.0** | **69.4** | 81.1 | 83.7 | **88.5** | 91.8 |
| Umbral-Poincaré-ball | 44.6 | 58.9 | 60.5 | 65.3 | 63.6 | 62.4 | 67.4 | 81.4 | 81.9 | 92.2 |
| Penumbral-half-space | 52.8 | **74.1** | 70.9 | 72.3 | 76.0 | 67.8 | **82.0** | 83.5 | 87.6 | 89.9 |
| Penumbral-Poincaré-ball | 44.6 | 60.8 | 62.7 | 68.4 | 67.9 | 60.8 | 69.5 | 78.2 | 84.4 | 92.6 |

F1 score (%) on WordNet noun, MCG, and Hearst with best numbers **bolded**

| Dataset | | Noun | | | | MCG | | | | Hearst | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Non-basic-edge Percentage | | 0% | 10% | 25% | 50% | 0% | 10% | 25% | 50% | 0% | 1% | 2% | 5% |
| Entailment Cone | d=5 | 29.2 | 78.1 | 84.6 | 92.1 | 25.3 | 56.1 | 52.1 | 60.2 | 22.6 | 45.2 | 54.6 | 55.7 |
| | d=10 | 32.1 | 82.9 | 91.0 | 95.2 | 25.5 | 58.9 | 55.5 | 63.8 | 23.7 | 46.6 | 54.9 | 58.2 |
| Umbral-half-space | d=5 | 45.2 | 87.8 | 94.2 | 96.4 | 36.8 | 80.9 | 85.0 | 89.1 | **32.8** | 63.4 | 77.1 | 80.7 |
| | d=10 | **52.2** | **89.4** | **95.7** | **97.0** | **40.1** | **81.9** | **87.5** | **91.3** | 32.6 | **65.1** | **81.2** | **86.9** |
| Penumbral-half-space | d=5 | 44.6 | 82.6 | 86.2 | 88.3 | 35.0 | 78.6 | 81.1 | 85.3 | 26.8 | 62.8 | 72.3 | 78.8 |
| | d=10 | 51.7 | 84.1 | 88.3 | 89.8 | 37.6 | 81.9 | 85.3 | 89.2 | 28.4 | 54.4 | 68.1 | 79.3 |

Half-space formulations of shadow cones achieve SOTA performance on edge prediction tasks across various data sets.