



山东师范大学
SHANDONG NORMAL UNIVERSITY

HyperMR: Hyperbolic Hypergraph Multihop Reasoning for Knowledge – based Visual Question Answering

Bin Wang, Fuyong Xu, Peiyu Liu* and Zhenfang Zhu *
School of Information Science and Engineering
Shandong Normal University

1 Motivation

2 Overall Framework

3 Hyperbolic Geometry

4 Hypergraph Construction

5 Hyperbolic Hypergraph Reasoning

6 Decoding

7 Main Experiments

8 Ablation Study

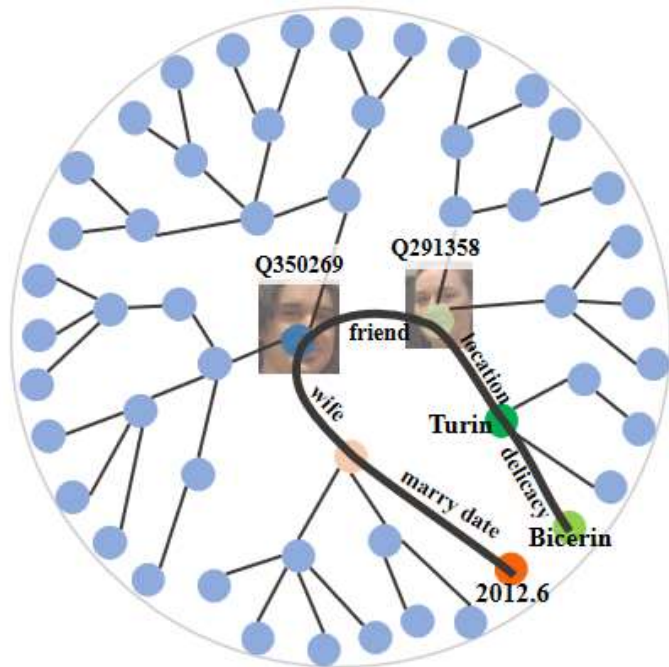


Fig.1 Knowledge embedding in Euclidean space. As the tree-like graph embedded in Euclidean space spreads, the edge nodes get closer to each other, leading to semantic distortion of the structure.

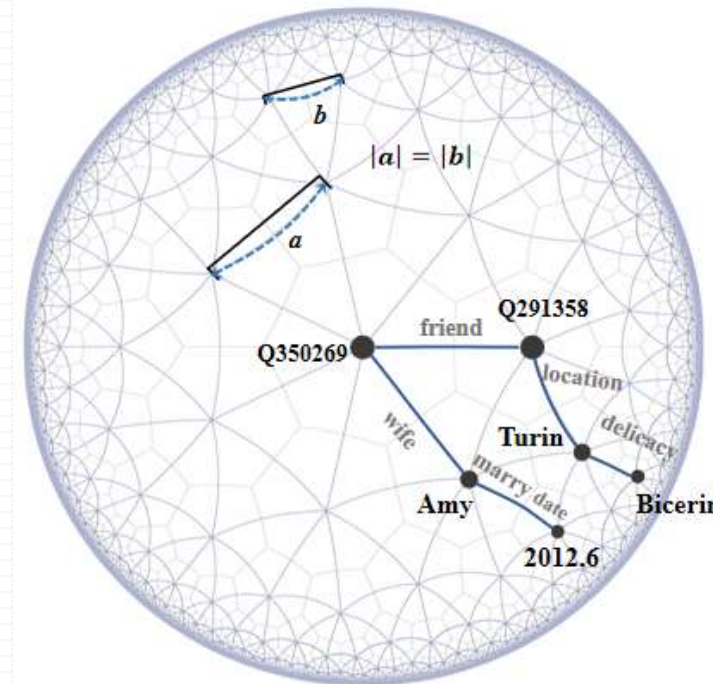


Fig.2 Knowledge embedding in Hyperbolic space. Each line segment in a Poincaré disc has the same length, e.g. a and b . Thus the relative positions of the nodes can be properly modeled.

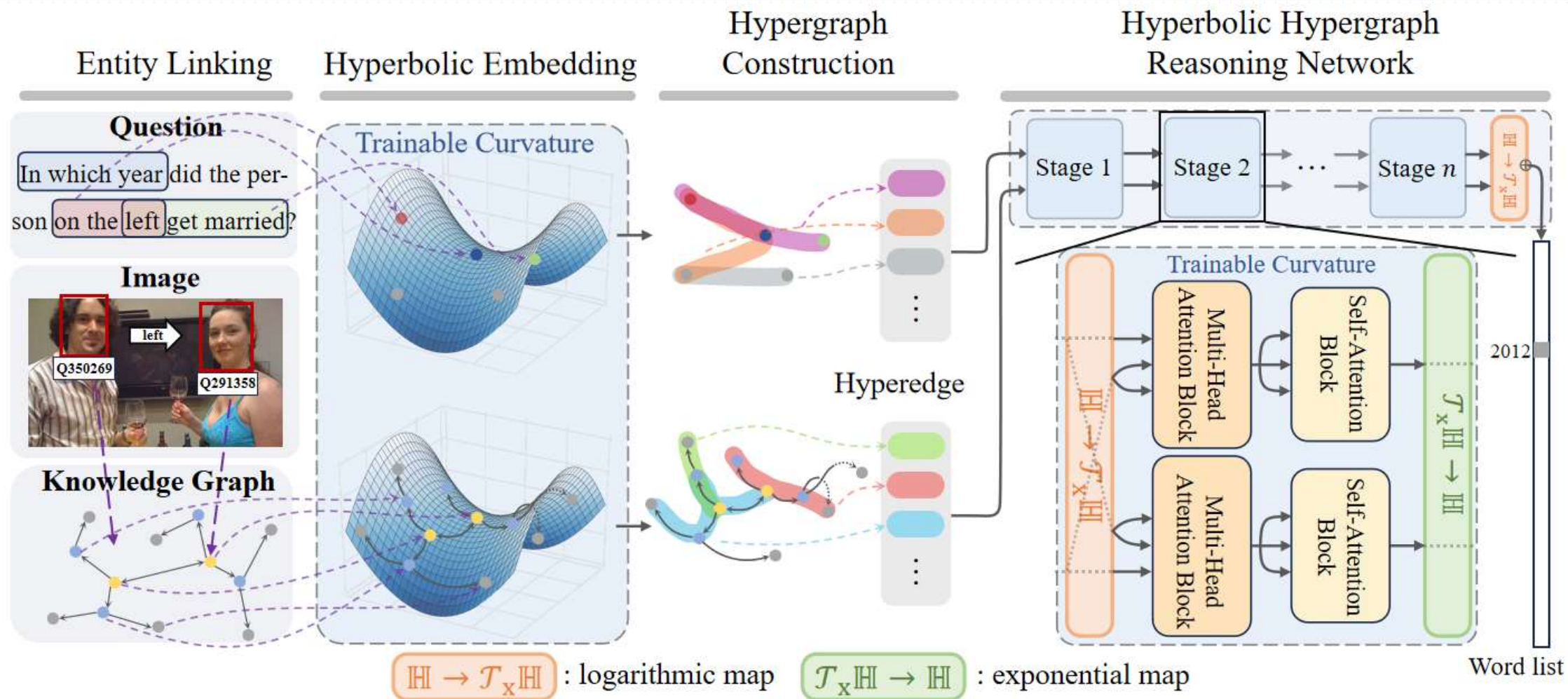
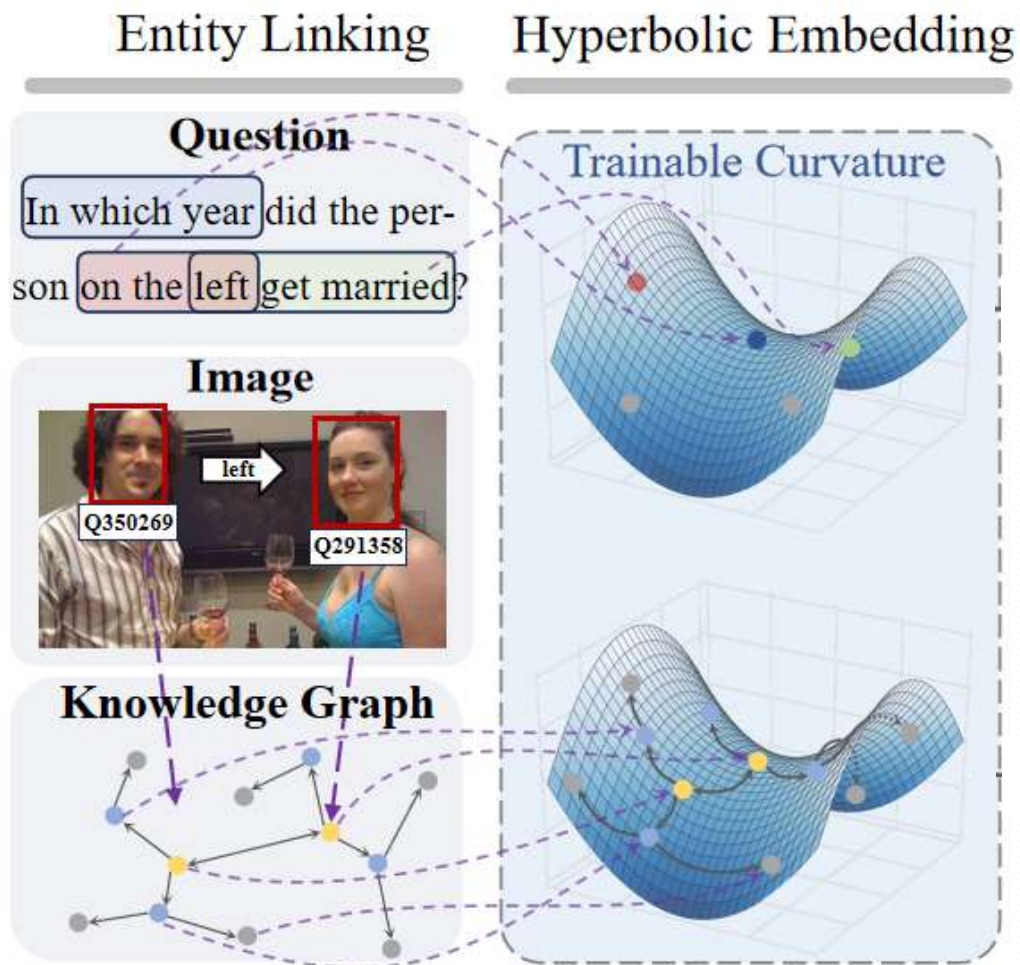


Fig.3 Model Framework

Definition:
$$\mathbb{H}^{n,c} = \left\{ x \in \mathbb{H}^n \mid \|x\|^2 < \frac{1}{c} \right\} \quad (1)$$

Möbius Addition:
$$x \oplus^c y = \frac{(1 + 2c \langle x, y \rangle + c \|y\|^2) x + (1 - c \|x\|^2) y}{1 + 2c \langle x, y \rangle + c^2 \|x\|^2 \|y\|^2} \quad (2)$$
 ^[1]

Poincaré Ball Distance:
$$d_c(x, y) = \frac{2}{\sqrt{c}} \operatorname{arctanh}(\sqrt{c} \|-x \oplus^c y\|) \quad (3)$$



Spatial Transformation:^[2]

$$\exp_x^c(z) = z \oplus^c \left(\tanh \left(\sqrt{c} \frac{\lambda_x^c \|z\|}{2} \right) \frac{z}{\sqrt{c} \|z\|} \right) \quad (4)$$

$$\log_x^c(y) = \frac{2}{\sqrt{c} \lambda_x^c} \tanh^{-1} \left(\sqrt{c} \|-x \oplus^c y\| \right) \frac{-x \oplus^c y}{\sqrt{c} \|-x \oplus^c y\|} \quad (5)$$

$$\mathbf{x}^{0, \mathcal{H}} = \exp_0^c \left((0, \mathbf{x}^{0, \mathcal{R}}) \right) = \left(0, \tanh \left(\sqrt{c} \|\mathbf{x}\| \right) \frac{\mathbf{x}}{\sqrt{c} \|\mathbf{x}\|} \right) \quad (6)$$

Fig.4 Entity Linking and Hyperbolic Embedding

Three words
to a lexeme

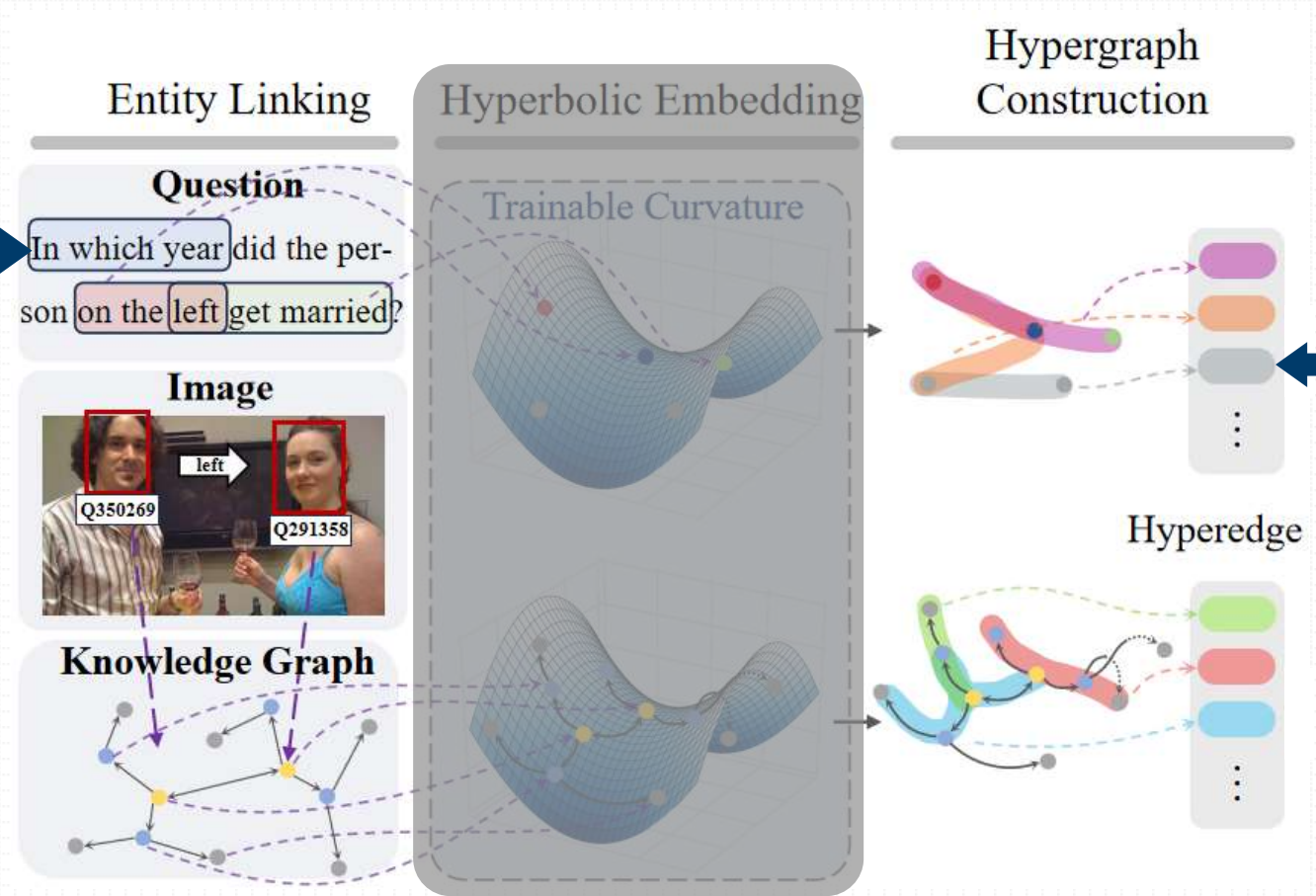
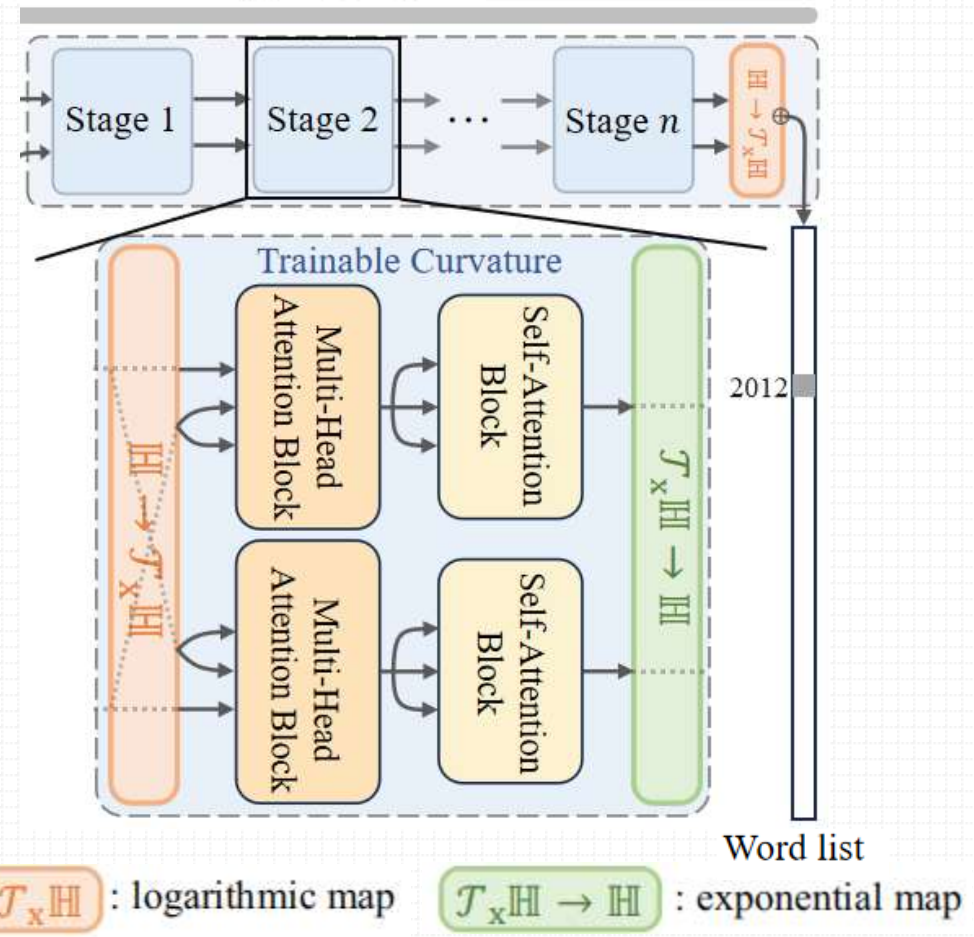


Fig.5 Hypergraph Construction

Hyperbolic Hypergraph Reasoning Network



Detail of Transformer:

$$Q_k = E^k W_{Q_k}; K_q = E^q W_{K_q}; V_q = E^q W_{V_q}, \quad (7)$$

$$Q_q = E^q W_{Q_q}; K_k = E^k W_{K_k}; V_k = E^k W_{V_k}, \quad (8)$$

where the multiplication in hyperbolic space:

$$W \otimes^c \mathbf{x} = \exp_0^c (W \log_0^c (\mathbf{x})). \quad (9)$$

Fig.6 Hypergraph Construction

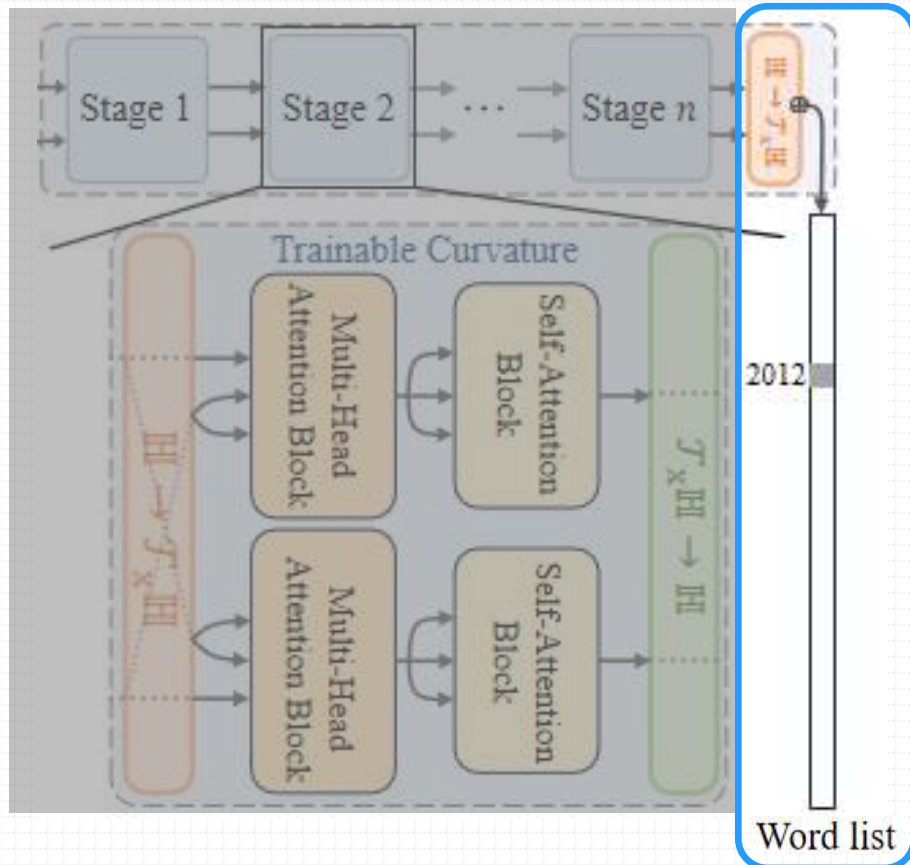


Fig.7 Decoding

Logarithmic mapping:

$$\log_x^c(y) = \frac{2}{\sqrt{c}\lambda_x^c} \tanh^{-1}(\sqrt{c}\| -x \oplus^c y \|) \frac{-x \oplus^c y}{\sqrt{c}\| -x \oplus^c y \|} \quad (10)$$

Probability of a word list:

$$p = j^{(0, \mathcal{R})} \mathbb{D}^T \quad (11)$$

Hyperbolicity $\delta=1.5$	Original (ORG)			Paraphrased (PRP)			Mean
	1-hop	2-hop	3-hop	1-hop	2-hop	3-hop	
BLSTM	-	-	-	-	-	-	51.0
GCN (Kipf and Welling, 2016)	65.7	67.4	66.9	65.8	67.5	67.0	66.7
GGNN (Li et al., 2016)	72.9	74.5	74.0	72.9	74.6	74.1	73.8
MeMNN (Sukhbaatar et al., 2015)	78.1	77.8	76.1	78.0	78.1	76.0	77.3
HAN (Kim et al., 2020)	77.5	77.5	77.2	77.1	77.4	76.9	77.3
BAN (Kim et al., 2018)	83.5	84.0	83.7	83.7	84.3	83.8	83.8
HyperTransformer (Heo et al., 2022)	88.1	90.2	91.0	87.8	90.5	90.7	89.7
DSAMR (Sun et al., 2024)	89.1	91.0	91.1	89.0	90.7	91.1	90.3
HyperMR	90.7	91.9	91.9	90.7	91.7	91.7	91.4

Tab.1 The accuracy of advanced methods for various types in Knowledge-aware VQA dataset.

Hyperbolicity δ	PathQuestion				PathQuestion-Large			
	PQ-2H	PQ-3H	PQ-M	Mean	PQL-2H	PQL-3H	PQL-M	Mean
	$\delta=2$	$\delta=2$	$\delta=2$		$\delta=0$	$\delta=1$	-	
Seq2Seq (Sutskever et al., 2014)	89.9	77.0	-	-	71.9	64.7	-	-
MemNN (Sukhbaatar et al., 2015)	89.5	79.2	86.8	85.2	61.2	53.6	55.8	56.9
KV-MemNN (Miller et al., 2016)	91.5	79.4	85.2	85.4	70.5	63.4	68.6	67.5
IRN (Zhou et al., 2018)	96.0	87.7	-	-	72.5	71.0	-	-
Embed (Bordes et al., 2014b)	78.7	48.3	-	-	42.5	22.5	-	-
Subgraph (Bordes et al., 2014a)	74.4	50.6	-	-	50.0	21.3	-	-
MINERVA (Das et al., 2018)	75.9	71.2	73.1	73.4	71.8	65.7	66.9	68.1
IRN-weak (Zhou et al., 2018)	91.9	83.3	85.8	87.0	63.0	61.8	62.4	62.4
SRN (Qiu et al., 2020)	96.3	89.2	89.3	91.6	78.6	77.5	78.3	78.1
HyperTransformer (Heo et al., 2022)	96.4	90.3	89.5	92.1	90.5	77.9	94.5	87.6
DSAMR (Sun et al., 2024)	98.4	91.1	91.7	93.7	95.6	81.7	98.8	92.0
HyperMR	96.2	90.5	89.5	92.1	96.3	83.7	98.9	93.0

Tab.2 Accuracy of advanced methods under PathQuestion and PathQuestion-Large datasets.

	PathQuestion			PathQuestion-Large			KVQA
	PQ-2H	PQ-3H	PQ-M	PQL-2H	PQL-3H	PQL-M	
# Entities	1,057	1,837	2,257	5,035	6,506	6,506	39,414
# Relations	14	14	14	364	412	412	18
# Knowledge facts	1,211	2,839	4,050	4,247	5,597	9,844	174,006
# Words	1,180	1,929	2,407	5,505	7,001	7,034	63,164
# QA pairs	1,908	5,198	7,106	1,594	1,031	2,625	183,007
# Answers	305	1,009	1,107	380	292	438	19,360

Tab.3 Statistics of benchmark datasets.

	Original (ORG)			Paraphrased (PRP)			Mean
	1-hop	2-hop	3-hop	1hop	2-hop	3-hop	
Single space w/ fixed curvature	90.2	91.2	91.5	90.1	91.2	91.4	90.9
Single space w/ trainable curvature	90.5	91.6	91.7	90.5	91.5	91.5	91.2
Multiple space w/ trainable curvature	90.7	91.9	91.9	90.7	91.7	91.7	91.4
HyperMR by hyperboloid	84.3	85.0	85.2	83.7	86.9	85.6	85.1

Tab.4 Ablation study results for the KVQA dataset at different settings.

	PathQuestion				PathQuestion-Large			
	PQ-2H	PQ-3H	PQ-M	Mean	PQL-2H	PQL-3H	PQL-M	Mean
Single space w/ fixed curvature	93.2	87.1	86.0	88.8	94.2	79.2	96.6	90.0
Single space w/ trainable curvature	94.2	87.7	90.1	90.1	95.1	81.3	97.5	91.3
Multiple space w/ trainable curvature	96.2	90.5	90.3	92.3	96.3	83.7	98.9	93.0
HyperMR by hyperboloid	84.8	81.4	82.1	82.8	94.4	78.9	95.4	89.6

Tab.5 Ablation study results for PQ and PQL datasets at different settings.



山东师范大学
SHANDONG NORMAL UNIVERSITY

HyperMR: Hyperbolic Hypergraph Multihop Reasoning for Knowledge – based Visual Question Answering

Bin Wang, Fuyong Xu, Peiyu Liu* and Zhenfang Zhu

Thanks