

Towards Robust Reasoning with Hyperbolic Representations

Akbar Karimi

Partner institutions:



Institutionally funded by:



Ministerium für
Kultur und Wissenschaft
des Landes Nordrhein-Westfalen



► ArithmAttack

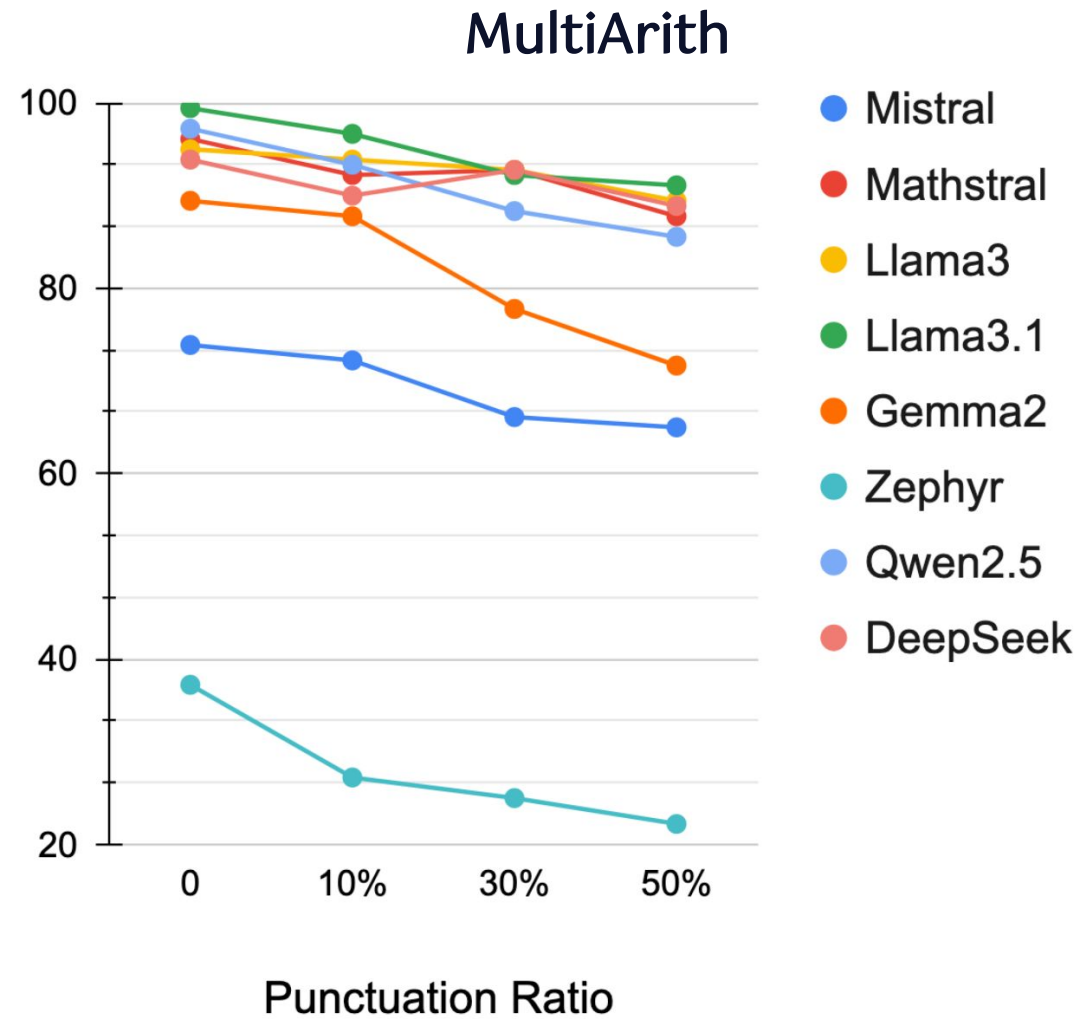
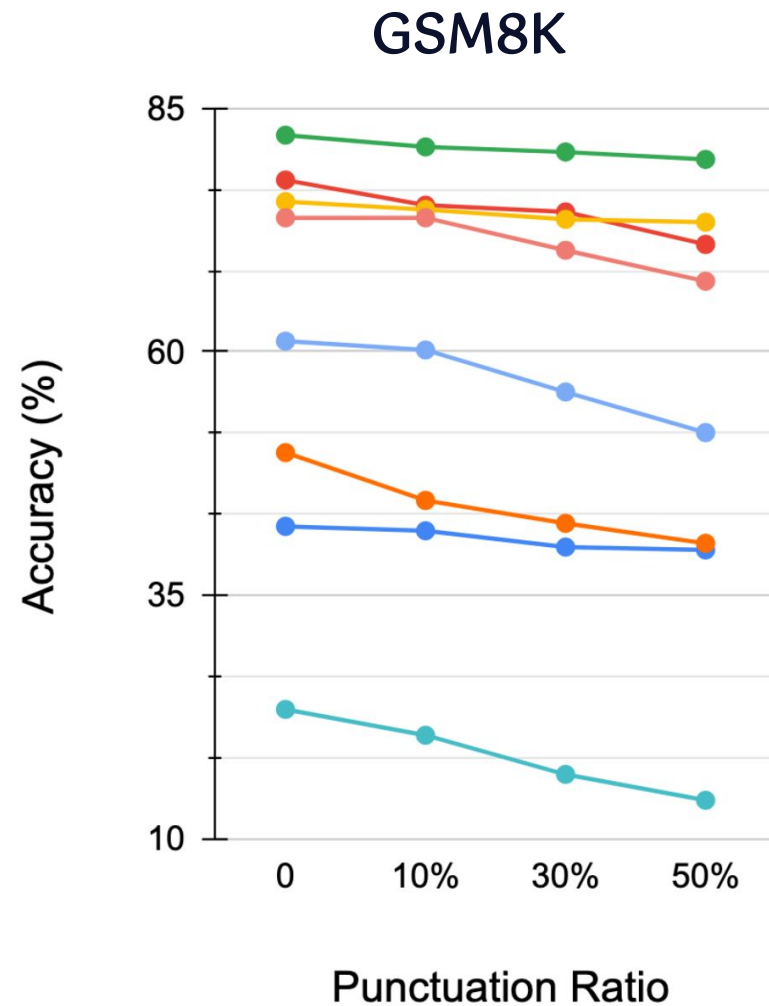
Arithmattack: LLM reasoning robustness to noise



Insert random punctuation marks into the inputs.

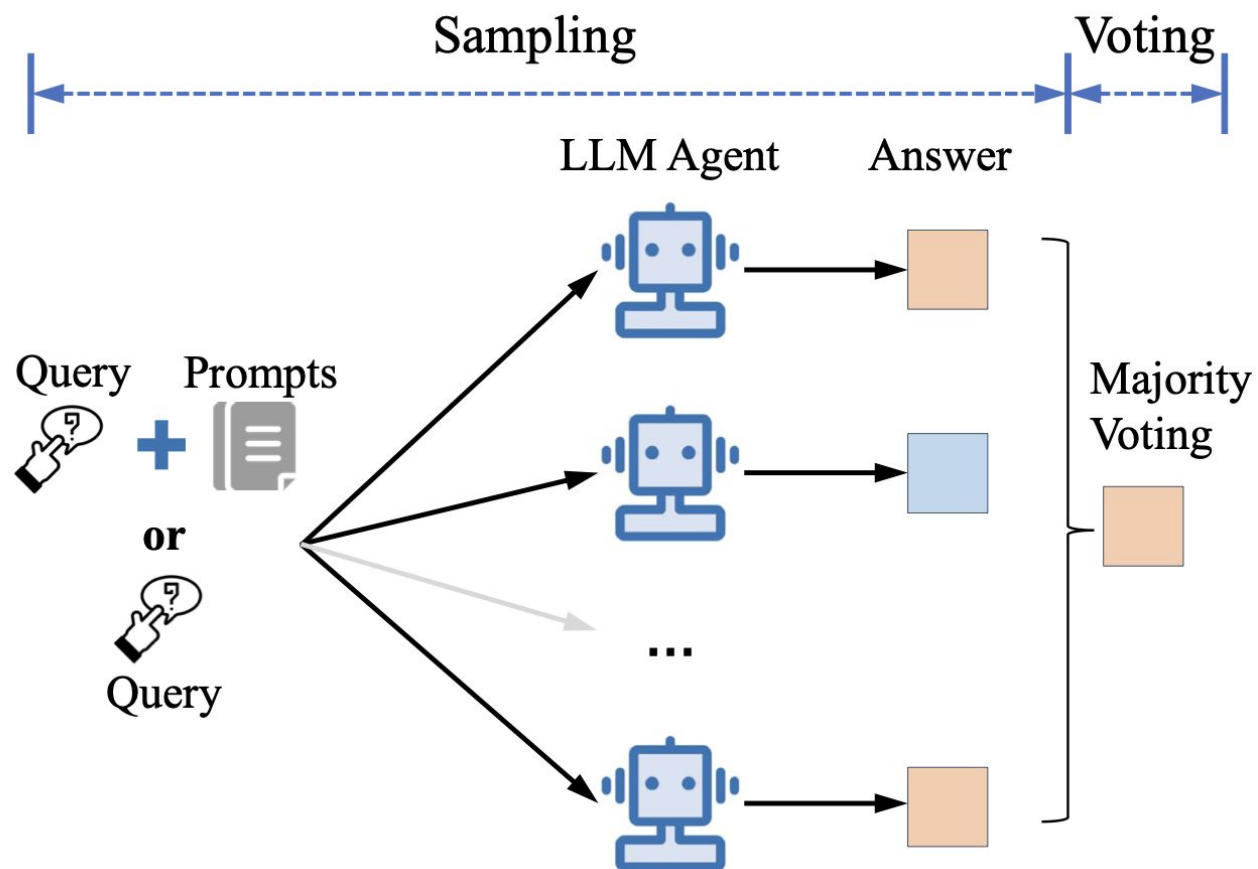
Real Prompt	Tiffany baked 8 brownies, but needed 17 total for her party. If she used 8 cups of flour on each one, how many cups of flour does she still need?	72 ✓	LLM Response
Noisy Prompt	? Tiffany baked 8 brownies, but needed 17 ! total for ; her party. If she : used 8 cups of flour on each one, how many cups of flour does she still need?	8 ✗	

Arithmattack: LLM reasoning robustness to noise



► Multi-agent Robustness

How Robust Is Multi-agent Reasoning to adversarial attacks in Math Problem Solving?

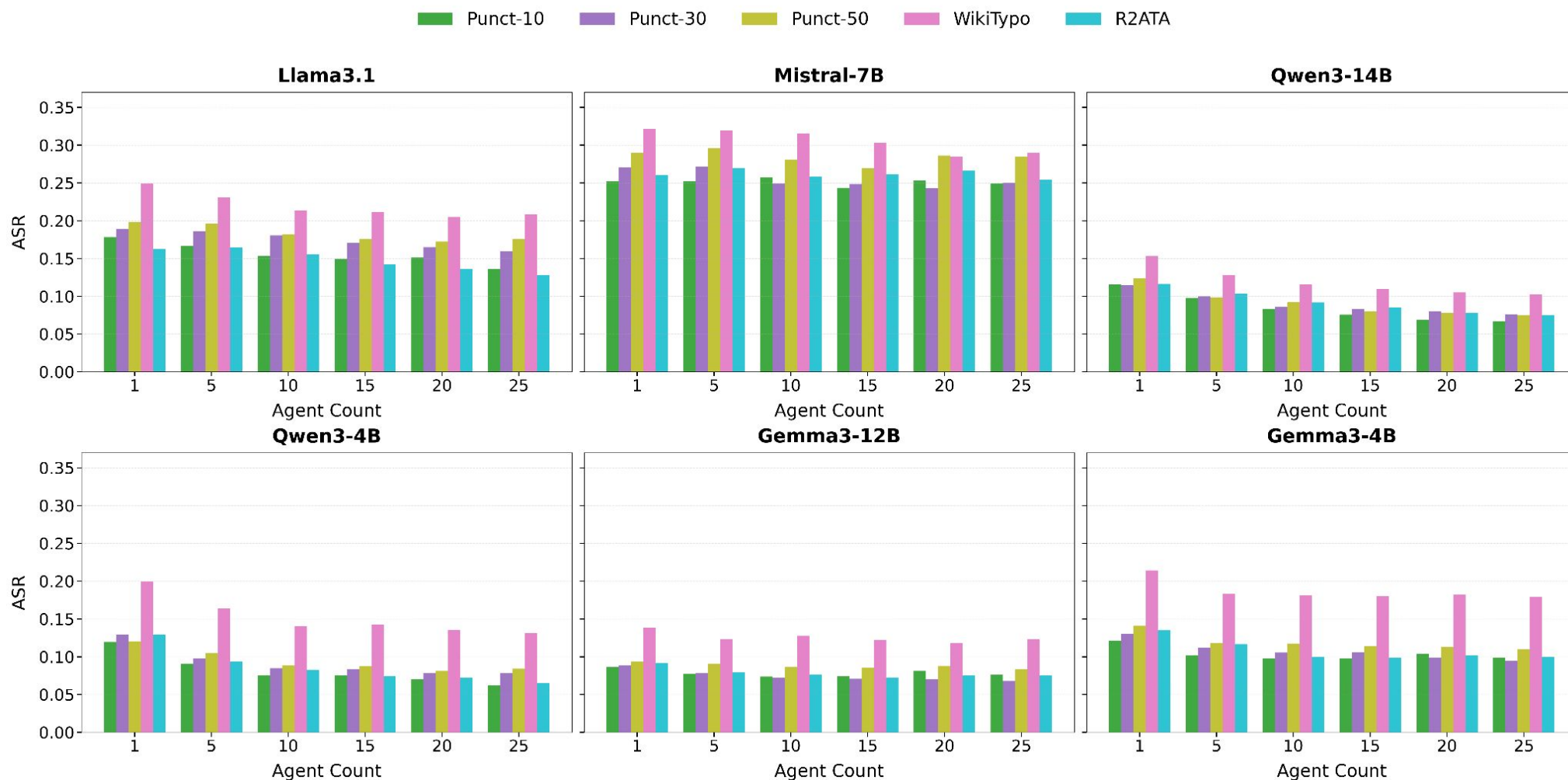


More Agents Is All You Need (Li et al. 2025)

How Robust Is Multi-agent Reasoning to adversarial attacks in Math Problem Solving?



ASR — Aggregated Across Datasets, Per Model



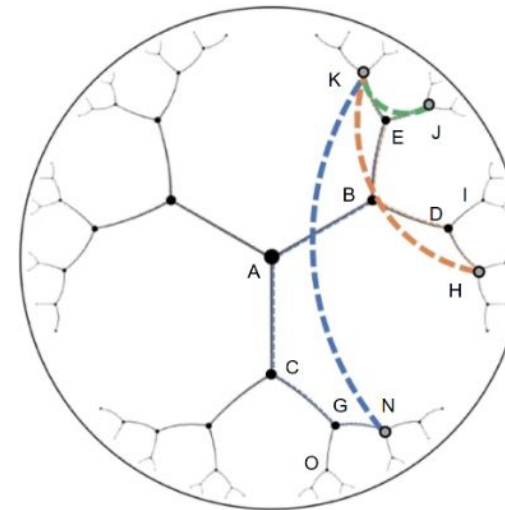
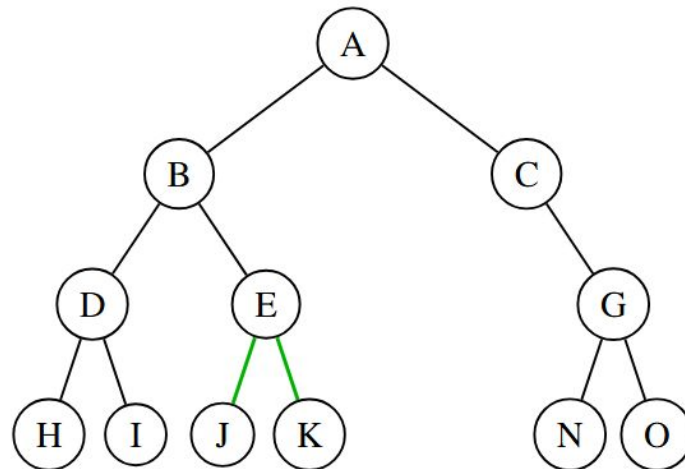
► Multi-hop Reasoning with Hyperbolic Representations

Why go hyperbolic?



Hyperbolic Models fit hierarchical structures naturally:

- ◆ Exponential Growth
- ◆ Distances modelled according to tree-structure



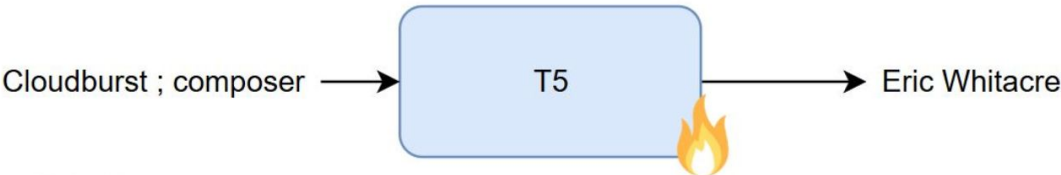
Multi-Hop Reasoning for Question Answering with Hyperbolic Representations



Question: Which country is the composer of song Cloudburst from?

Relevant Knowledge (Cloudburst, composer, Eric Whitacre)
(Eric Whitacre, country of citizenship, *American*)

Knowledge Integration



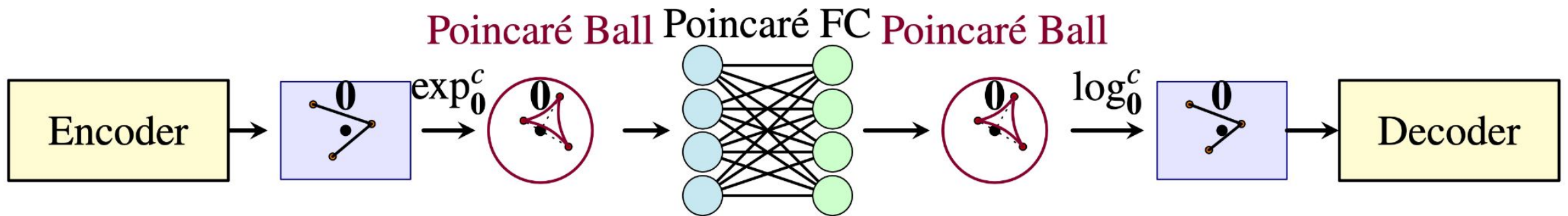
Parsing



Hopping



Architecture

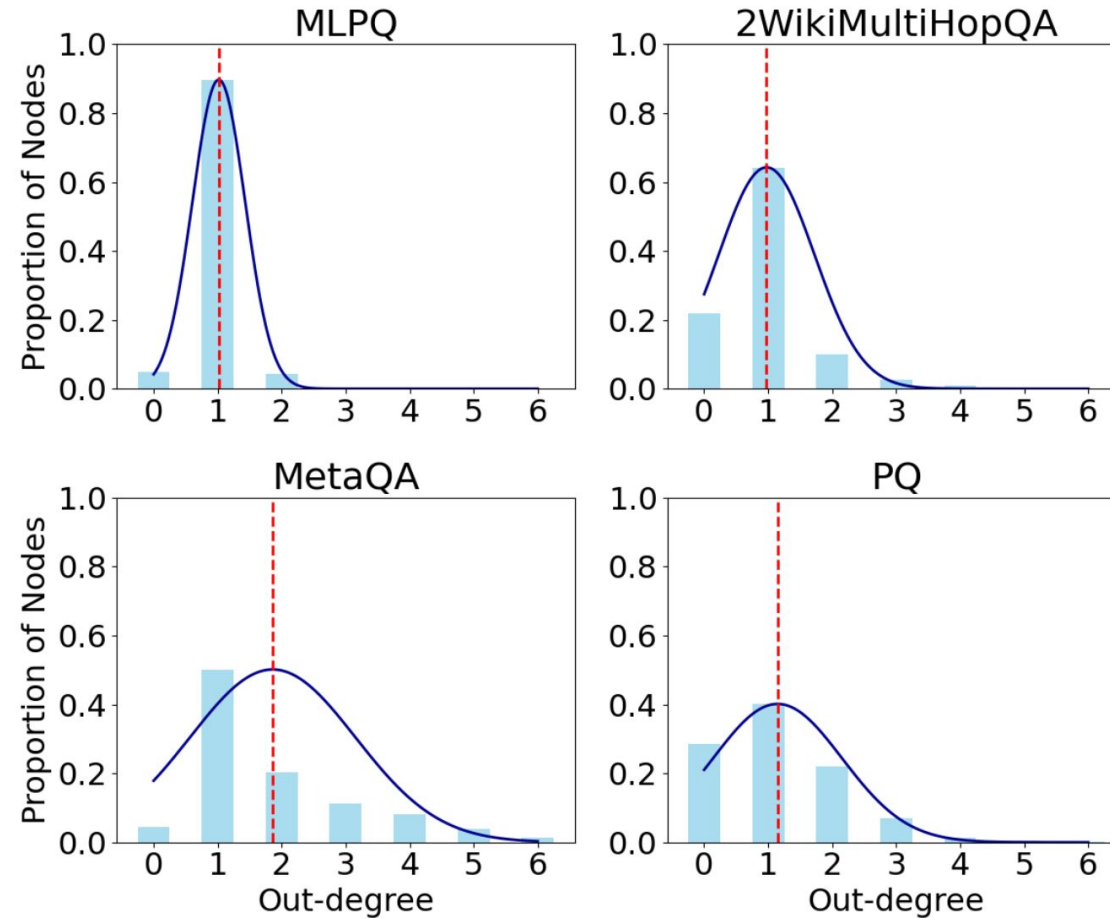


Main Results

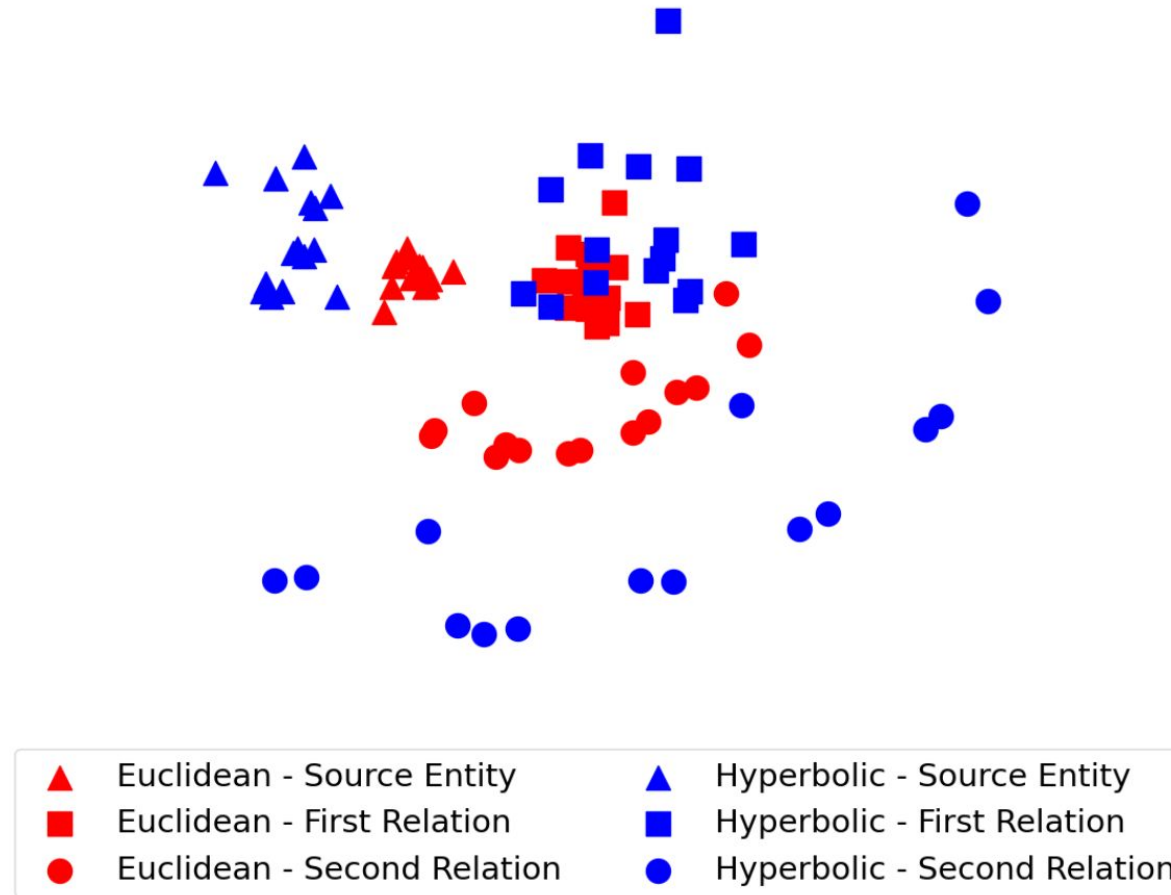


<i>Stage</i>	<i>Space</i>	2WikiHop	MetaQA	MLPQ	PQ
Hopping	Euclidean	44.36	22.92	81.03	18.28
	Hyperbolic	46.93	28.33	82.60	29.03
Parsing	Euclidean	88.60	95.51	97.08	100
	Hyperbolic	89.34	95.65	97.14	100

More hierarchy leads to more improvement.



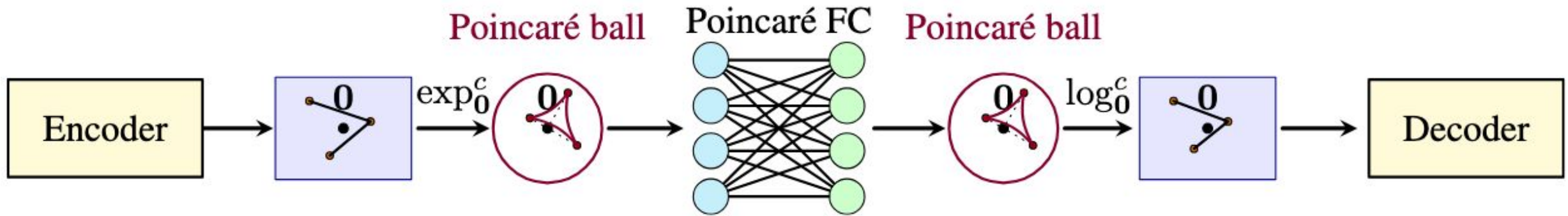
Hyperbolic embeddings are more spread out than Euclidean embeddings.



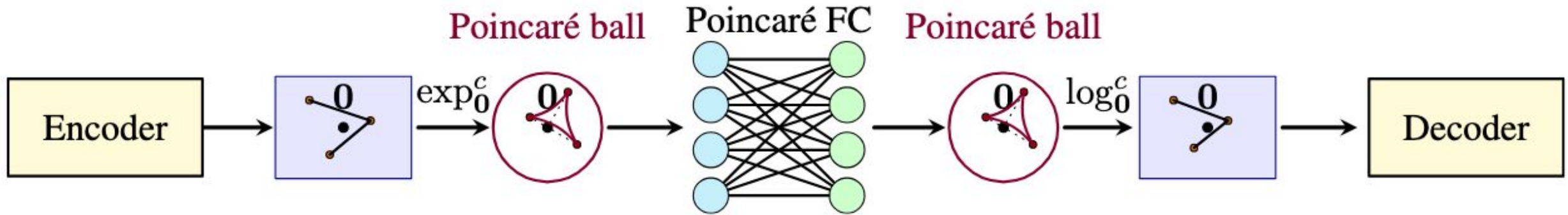


We should take the **data geometry** into account and consider using **hyperbolic representations** for hierarchical data.

And the future?

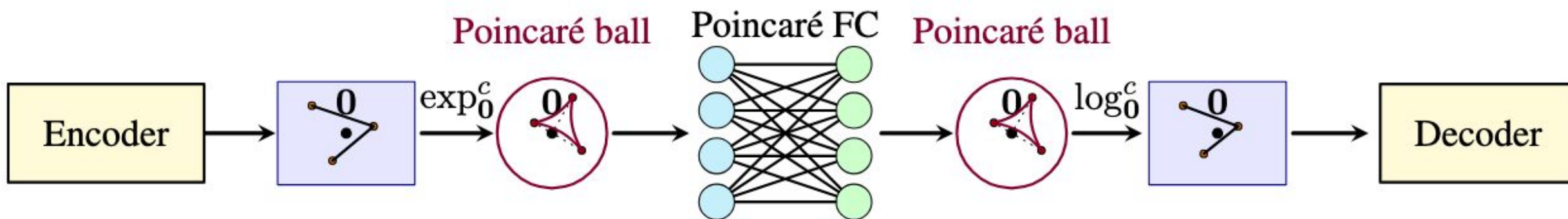


And the future?



Dataset	Knowledge Integration Additional	Random Walk Additional	Dev EM score
PQ	No additional	Linear	18.28
PQ	No additional	Hyperbolic	29.03

And the future?



Dataset	Knowledge Integration Additional	Random Walk Additional	Dev EM score
PQ	No additional	Linear	18.28
PQ	No additional	Hyperbolic	29.03
PQ	Hyperbolic (c=0.1)	Hyperbolic	43.62
PQ	Linear	Linear	11.17

Let's get in touch



ak@bit.uni-bonn.de
akkarimi.github.io