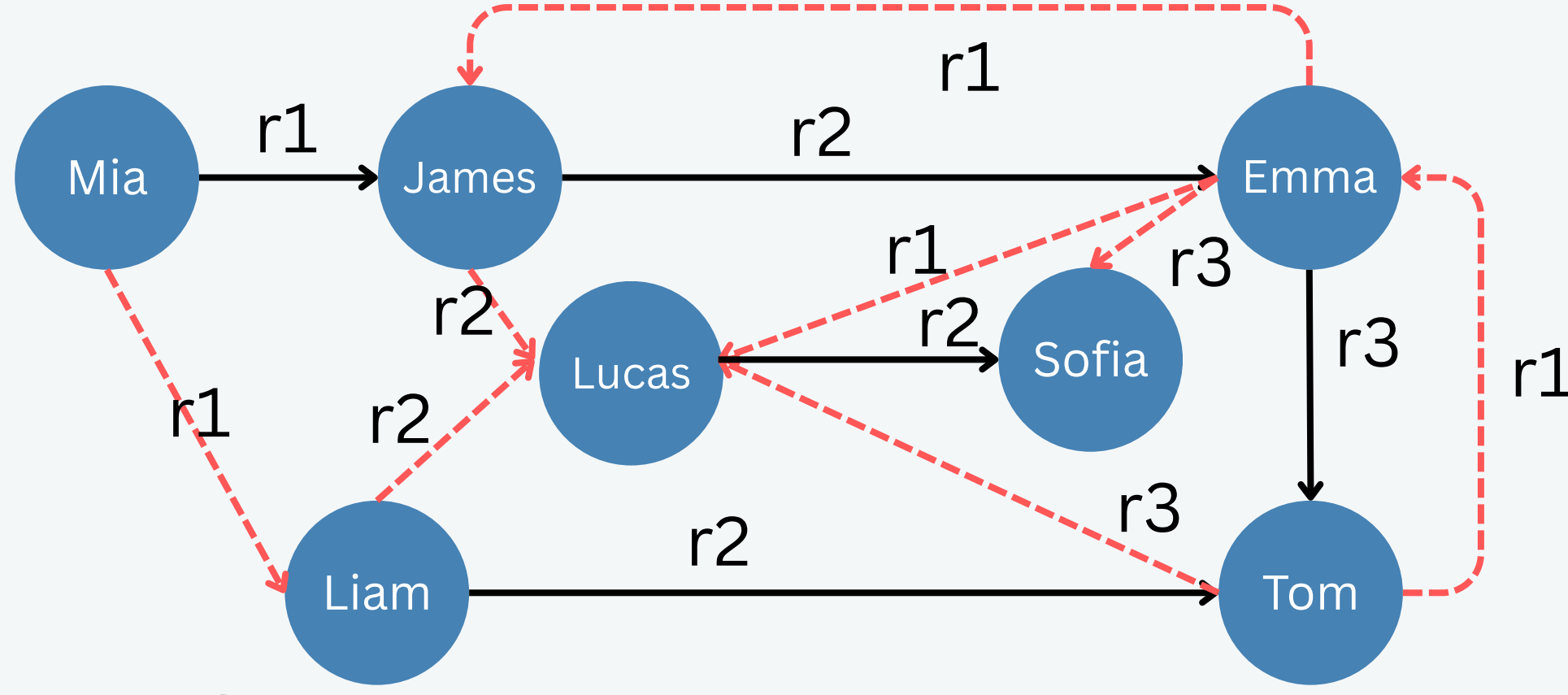


Querying graph with link prediction

- How to query an incomplete graph ?



Query format : (RegExp, set of start nodes)

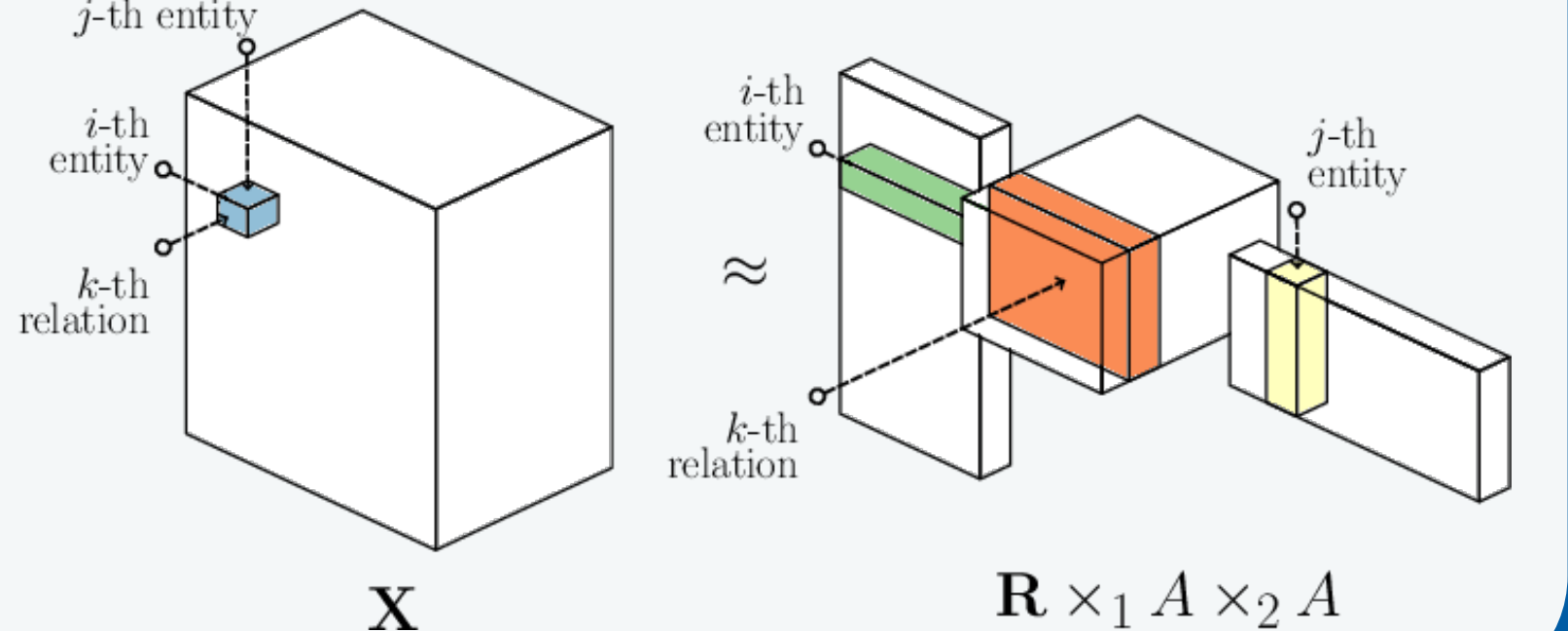
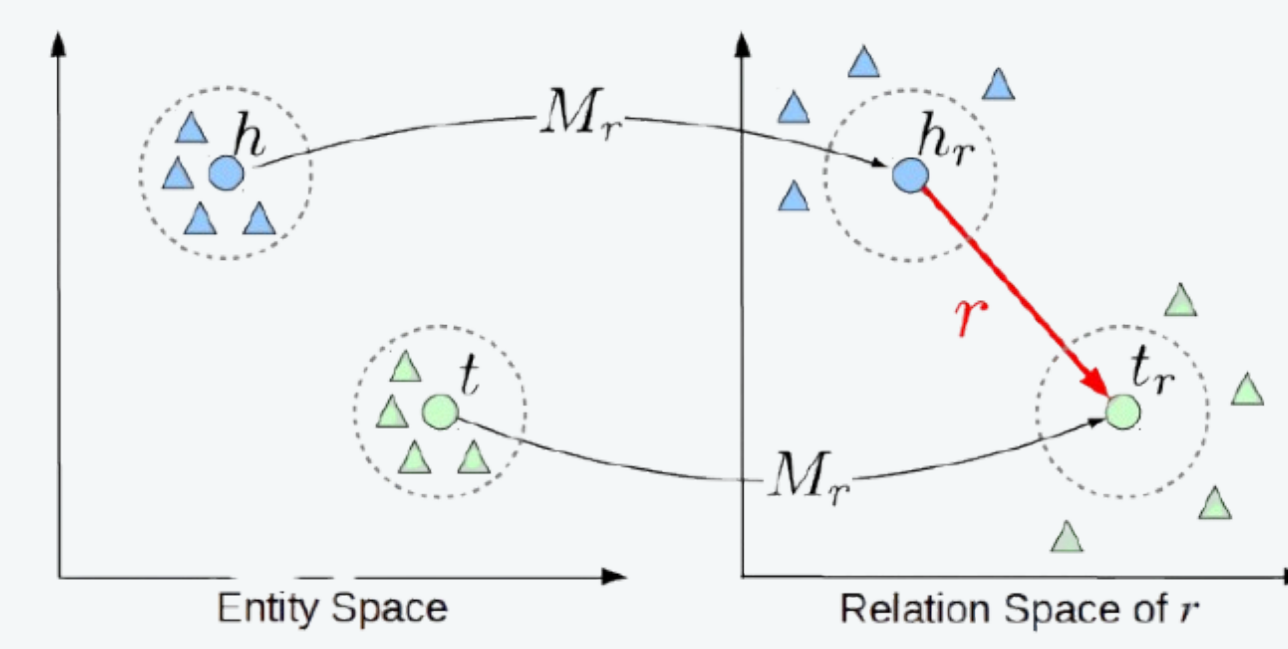
Query Example : ($r_1 . r_2 . r_3$, {Mia})

Graph Embedding

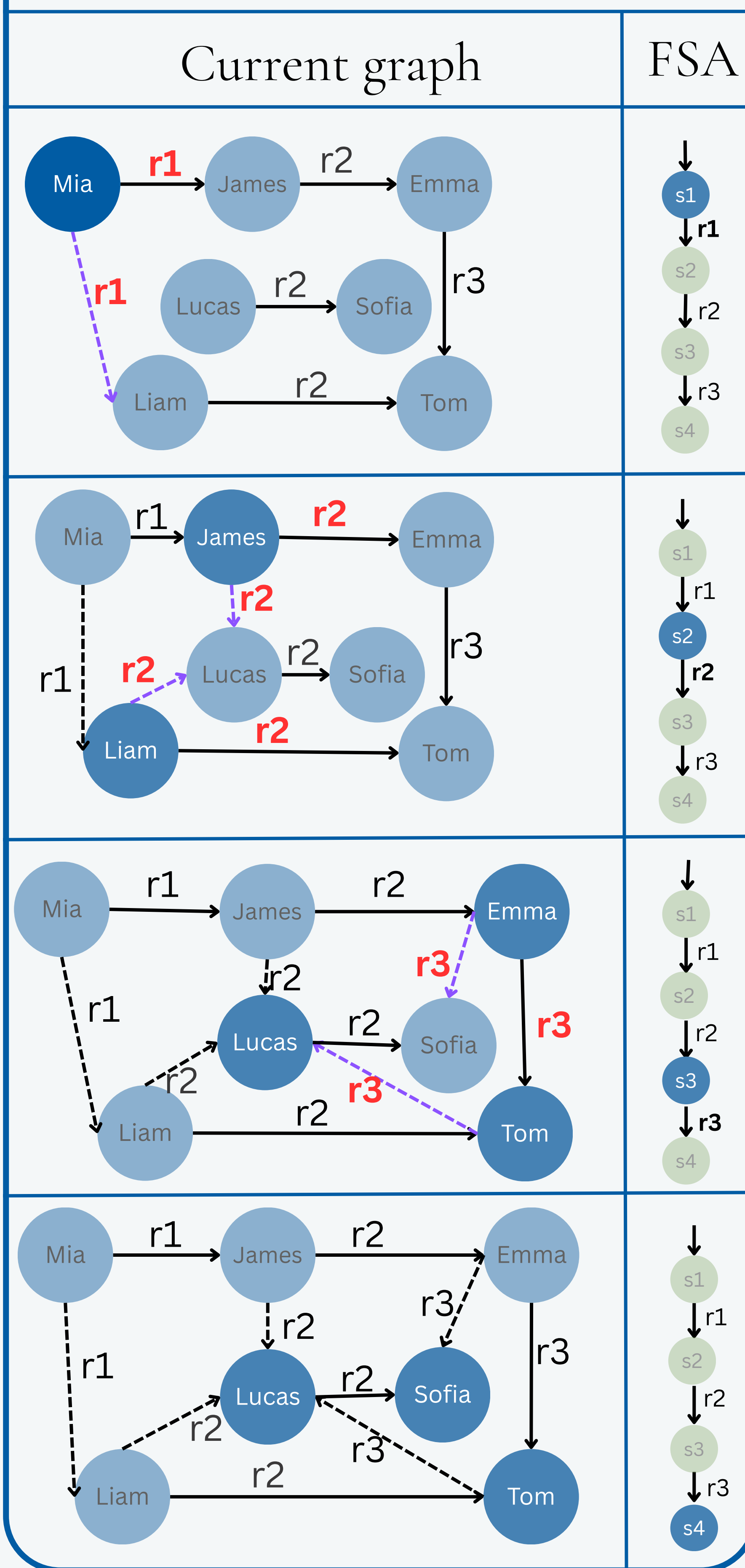
The function of a triple (h, r, t), where h is the head entity, r the relation, and t the tail entity, is defined as follows :

$$\text{TransR} : f_r(h, t) = ||h_r + r - t_r||_{L1/L2}$$

$$\text{RESCAL} : f_r(h, t) = h^T r t$$



Lazy - Prediction Querying



Experimental Results

Dataset	#Ent	#Rel	#Train	#Test	#Valid
Kinships	104	25	8544	1068	1070
UMLS	135	46	5216	652	661

Query types :

- size of 1, 2, 3
- with/without star
- ($r_1 . r_2 * . r_3$, {Mia})

G1, G2, G3 : Deleting the answers from the initial graph by 100%, 50%, 20%.

4 algorithm variations : cluster(yes/no), max number of prediction(yes/no).

5 strategies for prediction choices : rand, rrand, rscore, topk, topk+.

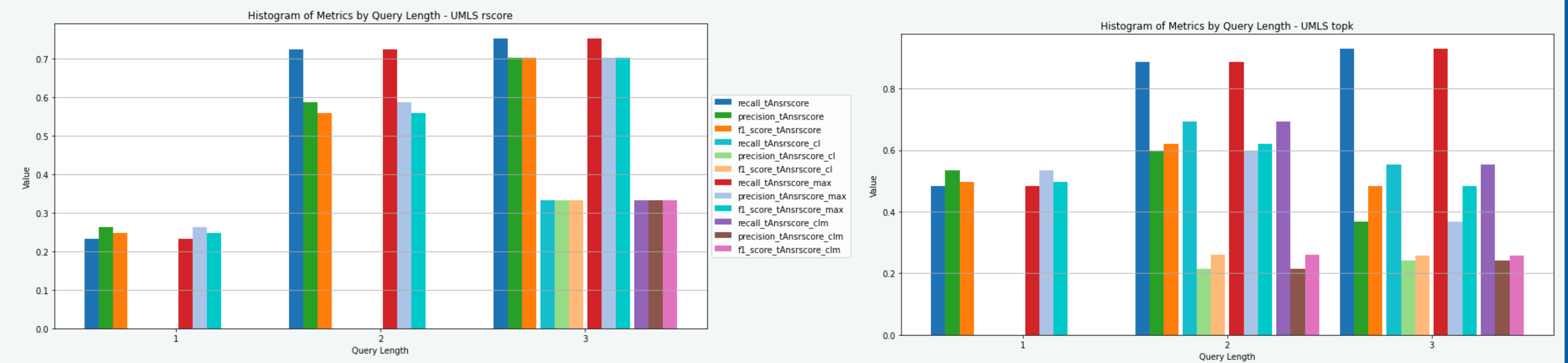


Figure 1 : Metrics by Query Length on UMLS G1 with TransR

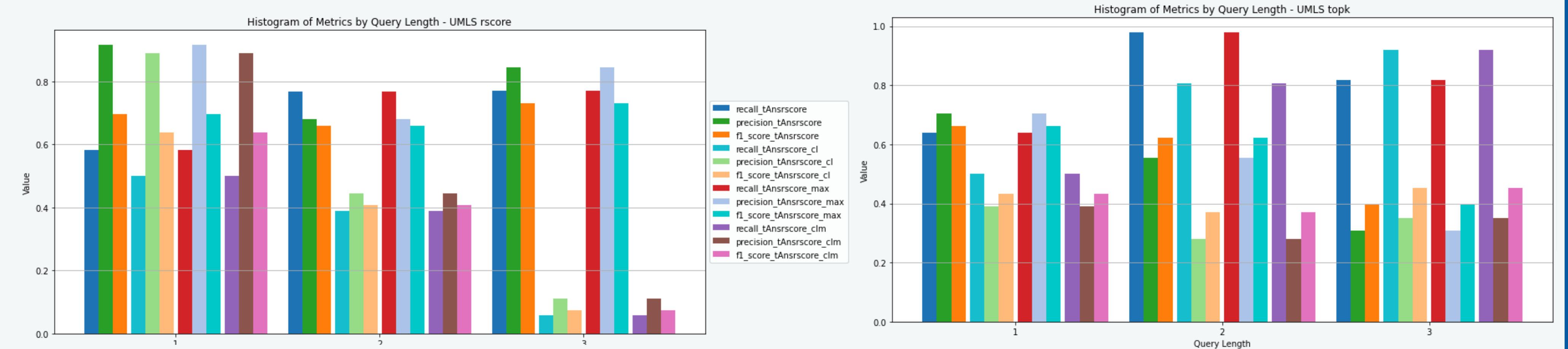


Figure 2 : Metrics by Query Length on UMLS G2 with TransR

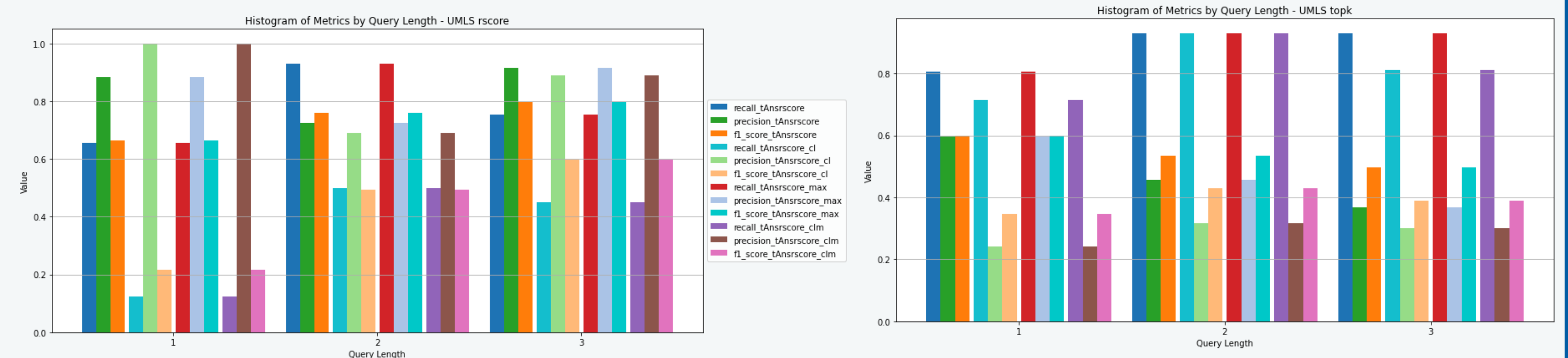


Figure 3 : Metrics by Query Length on UMLS G1* with TransR

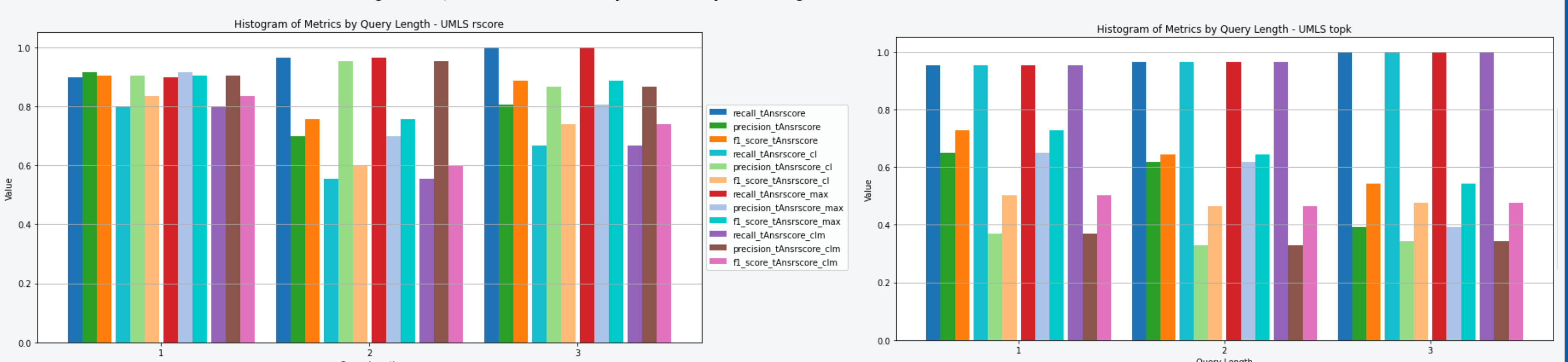


Figure 4 : Metrics by Query Length on UMLS G2* with TransR

Prediction choices

