

Data preprocessing for TransRHS

All original data could be found in data.rar:

../data/entity2id.txt
../data/relation2id.txt
../data/subRelation2relation.txt

The datasets are extracted from NELL, and the relation hierarchical structure we use is as follows:

e.g. (capitalCityOfCountry, subRelationOf, cityLocatedInCountry)

capitalCityOfCountry is the subRelation, while cityLocatedInCountry is the relation.

../data/entity2id.txt
entity list, the format is:
[Name id]

../data/relation2id.txt
relation list, the format is:
[Name id]

../data/subRelation2relation.txt
generalization relationship between relations, the format is:
[subRelation relation]