

Task Definition for Entity Recognition and Linking in Search Queries at NLPCC 2015

1. Overview

The main goal for this Entity Recognition and Linking Challenge at NLPCC 2015 is to evaluate the current advances of techniques in aligning named entities in short search queries to entities in a reference Chinese knowledge base.

2. The Task

In this Entity Recognition and Linking task, each query consists of an ID and a short search query. A system is expected to provide a reasonable linking result for the given query, where every named entity in this query should be linked into the knowledge base, if exists. For example, given a query:“射雕英雄传刘亦菲版”, a system is expected to link “射雕英雄传”into a knowledge base entity: “pk:tv:射雕英雄传(2006 年电视剧)” and **simultaneously**, “刘亦菲”into a KB entity “pk:per:刘亦菲”.

Example:

team-tag	run-tag	q-id	group-id	entity mention	KB-id
TeamABC	ABC-1	q-1	1	射雕英雄传	pk:tv:射雕英雄传(2006 年电视剧)
TeamABC	ABC-1	q-1	1	刘亦菲	pk:per:刘亦菲

For ambiguous queries, a system should enumerate at most 5 possible groups of interpretations for the query. For example, given a query: “射雕英雄传”, a system is expected to provide the following linking results: “pk:bk:射雕英雄传”, “pk:tv:射雕英雄传(1983 年电视剧)”, “pk:tv:射雕英雄传(1994 年电视剧)”, “pk:tv:射雕英雄传(2003 年电视剧)”, “pk:tv:射雕英雄传(2006 年电视剧)”, “pk:tv:射雕英雄传(2008 年电视剧)” and so on (at most 5 groups).

Example:

team-tag	run-tag	q-id	group-id	entity mention	KB-id
TeamABC	ABC-1	q-1	1	射雕英雄传	pk:bk:射雕英雄传
TeamABC	ABC-1	q-1	2	射雕英雄传	pk:tv:射雕英雄传(1983 年电视剧)
TeamABC	ABC-1	q-1	3	射雕英雄传	pk:tv:射雕英雄传(1994 年电视剧)
TeamABC	ABC-1	q-1	4	射雕英雄传	pk:tv:射雕英雄传(2003 年电视剧)
.....				

3. Scoring Metric

As we can see, for a given query, an EL system may output more than one group of interpretations. We thus evaluate system performances using **average F-scores**. Given a query q , the output of a system S^* , containing $|S^*|$ different groups of interpretations, each of which includes linking results for the named entities appearing in the given query. We compute the precision, recall and F-score by comparing S^* with human annotation S . The final **average F-scores** is obtained by averaging over each query.