Overview of the TREC 2010 Entity Track

Krisztian Balog
NTNU, Trondheim, Norway
krisztian.balog@idi.ntnu.no

Pavel Serdyukov
Yandex, Russia
pavser@yandex-team.ru

Arjen P. de Vries
CWI, The Netherlands
arjen@acm.org

1 Introduction

The overall goal of the track is to perform entity-oriented search tasks on the World Wide Web. Many user information needs concern entities (people, organizations, locations, products, ...); these are better answered by returning specific objects instead of just any type of documents.

Defining entities on the Web is still an unsolved problem. We settled on representing entities by their homepages, under the assumption that any entity of interest would have at least one homepage. The homepage URL is used as unique identifier. In this scenario, entity ranking corresponds to the task of returning the homepages of entities of a given type, that are relevant to the user's information need (represented as natural language text). We have to also consider that many entity queries could have very large answer sets (e.g., "actors playing in hollywood movies"); extra problematic with corpora the size of ClueWeb. In 2009, we decided therefore that finding associations between entities would be a more challenging one (in terms of modeling) and also a more manageable one (from a test collection building perspective) than finding associations between entities and topics, and defined the *Related Entity Finding (REF)* task (Balog et al., 2010). Related entity finding requests a ranked list of entities (of a specified type) that engage in a given relationship with a given source entity. REF ran as a pilot in 2009 and is the track's main task in this year; the document collection has been enlarged to the English subset of ClueWeb. We intend to repeat the REF task at least one more time in 2011.

One observation from the 2009 edition of the track is that many of the proposed approaches build heavily on Wikipedia and use it as a "semantic backbone": considering Wikipedia a large repository of entity names and types. Our goal is however not to evaluate entity retrieval over Wikipedia (this task has already been looked at in INEX, and a test collection exists), nor to limit ourselves to the (mostly popular) entities that are present in Wikipedia. As of this year, we are therefore not accepting Wikipedia pages as entity homepages.

The issue of combining (noisy) textual material (the Web) with semi-structured data (like Wikipedia or slightly more structured data sources like IMDB) is however an interesting line of research. As many data sources, and in particular those being constructed as so-called Linked Open Data (LOD), are naturally organized around entities, it would be reasonable to examine this problem in the context of entity retrieval. To foster research in this direction, we introduced the new *Entity List Completion (ELC)* pilot task. ELC is motivated by the same user scenario as REF, but with the main difference that entities are represented by their URIs in a Semantic Web crawl (the Billion Triple Collection). In addition, a small number of example entities (defined by their URIs) are made available as part of the topic definition. Our goal is to turn this pilot task to an "official" task in 2011.

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
NOV 2010		**00-00-2010 to 00-00-2010**

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
Overview of the TREC 2010 Entity Track	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Norwegian University of Science and Technology (NTNU),Trondheim, Norway,**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES
Presented at the Nineteenth Text REtrieval Conference (TREC 2010) held in Gaithersburg, Maryland on 16-19 November 2010. The conference was co-sponsored by the National Institute of Standards and Technology (NIST), the Defense Advanced Research Projects Agency (DARPA), and the Advanced Research and Development Activity (ARDA).

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			
unclassified	**unclassified**	**unclassified**	**Same as Report (SAR)**	**13**	

In the remainder of the paper we discuss the REF and ELC tasks in detail, in Sections 2 and 3, respectively. We summarize our findings and outline future plans in Section 4.

2 Related Entity Finding

Related Entity Finding (REF) ran as the main task of the track. Based on the experience gained from last year's pilot edition of the REF task, we implemented the following changes to the 2009 setup: (i) the document collection is enlarged to ClueWeb English, (ii) Wikipedia pages do not receive special treatment anymore, (iii) supporting documents are not required, (iv) "location" is added to target entity types, and (v) primary homepages receive more credit.

2.1 Task

The Related Entity Finding (REF) task is formulated as follows:

> Given an *input entity*, by its name and homepage, the *type of the target entity*, as well as the *nature of their relation*, described in free text, *find related entities* that are of target type, standing in the required relation to the input entity.

2.1.1 Input

For each request (query) the following information is provided:

- Input entity, defined by its name and homepage
- Type of the target entity (person, organization, product, or location)[1]
- Narrative (describing the nature of the relation in free text)

An example topic is shown below:

```
<query>
   <num>23</num>
   <entity_name>The Kingston Trio</entity_name>
   <entity_URL>clueweb09-en0009-81-29533</entity_URL>
   <target_entity>organization</target_entity>
   <narrative>What recording companies now sell
   the Kingston Trio's songs?</narrative>
</query>
```

2.1.2 Output

For each query, participants could return up to 100 answers (homepages). For each answer entity a single homepage must be returned; optionally, the name of the entity may also be returned.

2.1.3 Data collection

The document collection is the English portion of ClueWeb, comprising of approximately 500 million pages.

[1] Note that the input entity does not need to be limited to these four types.

2.2 Topics and assessments

Both topic development and relevance assessments were performed by NIST. For the 2010 edition of the track 50 new REF topics have been created. Out of these 47 ended up being assessed (excluded topics are: #35, #46, and #59). Participants were also requested to submit results for the 20 queries from last year.

We differentiate between *primary* and *relevant* homepages of a given entity: (i) a primary homepage is devoted to and in control of the entity, and (ii) a relevant homepage is devoted to the entity, but is not in control of the entity. By definition (and, unlike last year), the Wikipedia page of a given entity is non-relevant. Pages that only mention the entity (but are not about the entity) are also considered non-relevant. News articles and blog posts, even if exclusively about the entity, are not accepted as entity homepages. Products are required to have a separate page under the manufacturer's site.

All runs were pooled down to 20 records. Entity homepages were judged on a three-point relevance scale: (0) non-relevant, (1) relevant, or (2) primary. Names were judged as (0) incorrect, (1) inexact, or (2) correct, for the page returned. If the page is not primary, the correctness of the name is immaterial for the task. Finally, primary homepages are grouped together; primary documents in the same class are equivalent, and correct names for them are all valid.

2.2.1 Qrels

In the qrels file, the fields are:

```
topic doc name rel class rel_name
```

Where `topic` denotes the topic ID (corresponds to the `num` field of the topic definition), `doc` is a ClueWeb document ID, `name` is the normalized name of the entity, `rel` is the judgment for the document (0, 1, or 2), `class` is a class number for the document, and `rel_name` is the judgment for the name.

2.2.2 Evaluation metrics

The main evaluation measure we use is NDCG@R; that is, the normalized discounted cumulative gain at rank R (the number of primaries and relevants for that topic) where primary homepages get gain 3 and relevant homepages get gain 1. We also report on R-Precision (precision at rank R), and Mean Average Precision, both computed over primary pages only.

Note that evaluation results are not computed using the standard `trec_eval` tool, but a script developed specifically for the 2010 edition of the REF task[2].

2.3 Runs and results

Each group was allowed to submit up to four runs. Fourteen groups submitted a total of 48 runs; of those, 29 were automatic runs. Eight groups submitted a total of 19 manual runs.

The best automatic and manual runs from each group are shown in Table 1 and Table 2, respectively, while Table 5 displays all submitted runs. The Kendall tau rank coefficients indicate very strong correlation between the rankings of participating systems using various metrics (0.92 for NDCG@R vs. MAP, 0.94 for NDCG@R vs. rPrec, and 0.94 for MAP vs. rPrec).

[2] http://trec.nist.gov/data/entity/10/eval-entity.pl

Table 1: Best automatic REF runs from each group ordered by NDCG@R. The columns of the table (from left to right) are: runID, group, type of the run (Automatic/Manual), whether the Wikipedia subcollection received a special treatment (Yes/No), whether any external resources were used (Yes/No), NDCG@R, MAP, R-Precision, number of relevant retrieved homepages, and number of primary retrieved homepages. Highest scores for each metric are in boldface.

Run	Group	Type	WP	Ext.	NDCG@R	MAP	rPrec	#rel	#pri
bitDSHPRun	BIT	A	N	N	**0.3694**	**0.2726**	**0.3075**	**150**	314
FduWimET4	FDWIM2010	A	N	Y	0.3420	0.2223	0.2837	140	**333**
KMR1PU	Purdue_IR	A	Y	Y	0.2485	0.1555	0.2099	91	246
SuppHome	NiCT	A	N	Y	0.1696	0.0953	0.1453	74	187
ICTNETRun1	ICTNET	A	N	Y	0.1611	0.0839	0.1305	95	173
UWAT2	UWaterlooEng	A	N	Y	0.1393	0.0722	0.1223	96	154
LearnDPI	LIA_UAPV	A	N	Y	0.0766	0.0305	0.0591	72	81
G16	HPI	A	N	Y	0.0745	0.0357	0.0539	27	71
UAbaselinkA	UAmsterdam	A	Y	N	0.0496	0.0185	0.0349	34	81
ilpsA500	UAms	A	N	Y	0.0460	0.0178	0.0325	35	88
YahooEnHP	PITTSIS	A	N	Y	0.0375	0.0118	0.0229	37	42
CARDENSMBLE	CARD_UALR	A	N	N	0.0084	0.0000	0.0003	20	1

2.4 Approaches

The following are descriptions of the approaches taken by the different groups. These paragraphs were contributed by participants and are meant to be a road map to their papers.

BIT BIT Entity Group employs a logical sitemap constructor to extract hierarchical structures in order to enrich the anchor text model for finding more relevant pages. Those hierarchical structures, such as menus, navigational bars or breadcrumbs, indicate the logical relations between pages in the same site and the concise summary of pages in some sense. Under the assumption that items in similar visual presentations are probable similar in nature and to be classified in a group, they discriminate extracted entities by their locations in DOM tree and prefers to multiple entities in tables and lists. Finally, they find homepages from multiple sources and rank the homepages by their confidences and existences in ClueWeb09a English part for each candidate entity. (Yang et al., 2011)

Table 2: Best manual REF runs from each group ordered by NDCG@R. Highest scores for each metric are in boldface.

Run	Group	Type	WP	Ext.	NDCG@R	MAP	rPrec	#rel	#pri
bitRFRun	BIT	M	N	Y	**0.3897**	**0.2876**	**0.3209**	**153**	**319**
FduWimET3	FDWIM2010	M	Y	Y	0.3376	0.2218	0.2886	116	297
KMR3PU	Purdue_IR	M	Y	Y	0.2917	0.1916	0.2505	93	296
EntityHP1	PITTSIS	M	N	Y	0.2884	0.1664	0.2258	140	278
PRIS2	PRIS	M	N	Y	0.2846	0.1607	0.2296	128	312
SIELRUN1	SIEL_IIITH	M	Y	Y	0.1576	0.1019	0.1414	38	198
ilpsM50agfil	UAms	M	N	Y	0.0718	0.0331	0.0496	36	99
UAcatslinkA	UAmsterdam	M	Y	N	0.0708	0.0485	0.0678	29	84

CARD_UALR To find relevant entities and their homepages, first, we identified the entities and their types using Stanford Named Entity Recognizer. Due to its limitations, we could only identify PERSON, LOCATION and ORGANIZATION type entities. Next, an entity-entity co-occurrence graph was established. If two entities co-occurred in a webpage more than a specified threshold, the two entities were linked. Given the query entity, relevant entities are extracted based on a novel centrality measure (Cumulative Structural Similarity-CSS) using the intuition that an important entity will share many common neighbors with adjacent entities. Additionally, PageRank, HITS and Ensemble-based approaches are submitted. (Agarwal et al., 2011)

FDWIM2010 The FDWIM group proposes a multiple-stage retrieval framework for the task of related entity finding. In the document retrieval stage, search engine is used to improving the retrieval accuracy. In the next stage, they extract entity with NER tools, Wikipedia and text pattern recognition. Then stoplist and other rules is employed to filtering entity. Deep mining of the authority pages is effective in this stage. In entity ranking stage, many factors including keywords from narrative, page rank, combined results of corpus-based association rules and search engine are considered. Finally, an improved feature-based algorithm is proposed for the entity homepage detection. (Wang et al., 2011a)

HPI The approach of the HPI-group studies in particular the exploitation of advanced features of different Web search engines to achieve high quality answers for the related entity finding task. Thus, the system preprocesses a topic using part-of-speech tagging and synonym dictionaries, and generates an enriched keyword query employing advanced features of the particular Web search engine. After retrieving a corpus of documents, the system constructs an extraction rule that consists of the source entity (and synonyms), the target entity type and words that should occur in the context of both (taken from the narrative relation description). After the extraction of potentially related entities, they are subjected to a deduplication mechanism and scored for each document with respect to the distance to the source entity. Finally, these scores are aggregated across the corpus by incorporating the rank position of a document. For homepage retrieval the HPI-system further employed advanced features of the used Web search engines - for instance to retrieve candidate URLs by queries such as "entity in anchor". Homepages are ranked by a weighted aggregation of feature vectors. The weight for each of the 17 used features was determined beforehand using a genetic learning algorithm. The submitted runs compare the performance of the three most popular search engines, that were employed by the system. (Hold et al., 2011)

ICTNET The ICTNET group proposes a bipartite graph reinforcement model for entity ranking. Firstly, the candidate entities are extracted from related text snippets and are ranked based on a probabilistic model. Secondly, the lists which may contain several target entities are also extracted. Thirdly, a bipartite graph is constructed in which candidate entities and lists are considered as the two disjoint sets of graph vertices. Finally, the reinforcement algorithm is applied over the graph to get the final score for each candidate entity. For the homepage finding, google is used to search for top-K urls and some heuristic rules are used to identify the real homepage. (Cao et al., 2011)

LIA_UAPV The LIA and iSmart group proposes a Question Answering approach to address REF. They were focused on a way to validate candidate named entities at the end of the QA process. For this, they proposed an unsupervised way to determine in what extent a named entity belongs to a given type. They started by extracting a fined grained type from topic's narrative field (e.g. "teammates"), collected web pages about it and computed word distribution on them. They used similar process for each candidate named entity. Then, they computed a degree of similarity between an entity and the type by comparing

their word distribution. Finally, they proposed four different ways to re-rank candidate named entities. (Bonnefoy et al., 2011)

NiCT In 2010, the NiCT group mainly focused on improving target entity extraction and entity ranking, both of them play vital roles in the REF system. A Named Entity Recognition tool is first used to extract entities that match types of target entities such as organization, person, etc. Secondly, dependency tree-based patterns learnt automatically are employed to filter out the extracted entities that do not match fine-grained types of name entities such as university, airline, author, etc. In ranking part, a dependency tree-based similarity approach is proposed, which is better than language model. (Wu et al., 2011)

PITTSIS Our method is based on a two-layer probability model for integrating document retrieval and entity extraction together. The document retrieval layer finds highly relevant documents, and the entity extraction layer extracts the right entities. Our goal in this year TREC is to set up a frame work for evaluating and exploring each individual layer as well as the overall workflows. This method helps to reduce the overall retrieval complexity while keeping high accuracy in locating target entities. (Li and He, 2011)

PRIS The PRIS group proposes Document-Centered Model (DCM) and Entity-Centered Model (ECM) for the entity finding task. In DCM, documents are seen as a bridge. Both probabilities of a query and entity with respect to a document are estimated. In ECM, snippets extracted from documents are at the bottom to support entities. BM25 method is also introduced into ECM besides indri retrieval model. Another improvement aims to entity extraction. Special web page, NER tool and entity list generated by some rules are all taken into account. (Wang et al., 2011b)

Purdue_IR In the related entity find (REF) task, we generally follow our previous work on TREC Entity 2009. The structures of tables and lists are further investigated to extract related target entities from them. Moreover, we infer the type of target entities from the query topic and infer the type of candidate entities from their profiles, and then match the two types. (Fang et al., 2011)

SIEL_IIITH We use external resources like Wikipedia and Web, as Clueweb Category A dataset is not available. We extract all entities from Wikipedia using pattern finding techniques and indexed them with their type. We searched query in this index to find target entities. We use web search to find target entities not present in Wikipedia index. We then combine both the results to get final ranking. We then used Clueweb's URL-DocId mapping to find urls of target entities present in Clueweb dataset and present corresponding DocID as final results. This approach give satisfactory results in the absence of Clueweb dataset. (Shaik et al., 2011)

UAms To address REF we look for homepages of entities of the target type that co-occur with the source entity in contexts of a certain size, emphasizing contexts that contain terms from the relation (the narrative provided with a topic). We experimented with context size by varying a window size parameter. To perform filtering based on type and homepage finding we use Freebase, which provides category labels and homepage URLs. To remove NER errors we restrict the candidate entities to those in Freebase. In addition to Freebase homepage URLs we submitted entity strings to a web search engine to find homepages. (Bron et al., 2011)

UAmsterdam The University of Amsterdam, group of Jaap Kamps, participates only in the main related entity finding task, and uses Wikipedia as a pivot to search for entities. The

approach is very similar to last year's approach. Wikipedia topic categories are manually assigned to the query topics, which are more specific as the given target categories. These more specific target categories are used to retrieve entities within Wikipedia. To search web entities the external links in Wikipedia are used, and an anchor text index is searched. (Kamps et al., 2011)

UWaterlooEng The University of Waterloo investigated whether related entity finding problem can be addressed by unsupervised approaches that rely primarily on statistical methods and common linguistic tools, such as named-entity taggers and syntactic parsers. An initial candidate list of entities is extracted from top ranked documents retrieved for the query, and then refined using a number of statistical and linguistic methods. One of the key components of their method consists of finding hyponyms of the category name specified in the narrative, representing candidate entities and hyponyms as vectors of grammatical dependency triples, and calculating similarity between them. (Vechtomova, 2011)

2.5 Common themes

In this subsection we discuss some general tendencies that we observed among participating systems this year.

2.5.1 Manual runs

The fraction of manual runs, as opposed to automatic ones, was relatively high this year (19 out of 48 runs); two teams (PRIS, SIEL_IIITH) actually submitted manual runs only. Here, we briefly review the various types of interventions in the retrieval process that groups resorted to in their manual runs.

The FDWIM team constructed queries for retrieval of support documents manually. The PRIS group checked the correctness of extraction for some part of entities and boosted the score of manually recognized entities. The Purdue_IR group submitted a manual run in which the types of target entities were chosen manually. On a similar account, the UAmsterdam team assigned more specific entity types to each query by hand. The UAms group did not interfere much with the automatic execution of the retrieval workflow; they merely removed stop words and added the base forms of verbs and singular forms of plural terms to the narrative manually.

2.5.2 External resources and Wikipedia

Another observation we make is that most runs (39 out of 48) used external resources. This is much higher than in last year (15 out of 41). On the other hand, the reliance on Wikipedia has decreased slightly (14 out of 48 runs treated Wikipedia in a special way, in contrast to last year's 16 out of 41). The former, in part, might be necessitated by the move to ClueWeb English; groups that could not handle the collection resorted to Web search engine APIs. The latter is probably due to the fact that Wikipedia pages are no longer accepted as relevant answers.

The BIT groups uses Google and Realnames search engines for homepage finding. HPI queries Freebase to find synonyms of entities; these, then, are used to construct a query which is sent to Google, Bing, or Yahoo!. Moreover, they make extensive use of search operators when querying Google. LIA_UAPV uses the Yahoo! search engine to find the canonical form of a person name and then to find support documents (again, by querying Yahoo!). Finally, they use Yahoo! to find the homepages of retrieved entities. UWaterlooEng, PITTSIS, and NiCT also use Yahoo! to find support documents. NiCT uses YAGO/DBPedia data to learn patterns for "isA" relations. ICTNET uses Wordnet to find synonyms. UAms uses Bing, as well as Freebase/DBPedia.

2.5.3 Named entity recognition

Based on the participating systems' descriptions it seems that only the UAmsterdam group did not use named entity tagging. Most teams (BIT, CARD_UALR, FDWIM2010, ICTNET, LIA_UAPV, PRIS, Purdue_IR, and UAms) used the Stanford Named Entity Recognizer or some extension of it. HPI employed the SAP Business Objects Thingfinder, NiCT used the UIUC NE toolkit, and UWaterlooEng applied an LBJ-based Named Entity Recognizer.

3 Entity List Completion

The Entity List Completion (ELC) task has been introduced this year and ran as a pilot.

3.1 Task

ELC addresses essentially the same task as REF does: finding entities that are engaged in a specific relation with an input entity. There are three differences to REF:

- Entities are not represented by their homepages, but by a unique URI (from a specific collection, a sample of the Linked Open Data cloud),
- A small number of known relevant entities are made available as part of the topic definition, as examples.
- The target type is mapped to the most specific class within the DBPedia ontology.

3.2 Data collection

We use the Billion Triple Challenge (BTC) collection[3], a publicly available Semantic Web crawl; we consider this collection as a reasonable sample of Linked Open Data (LOD). Not all nodes in this Semantic Web graph are entities; identifying the nodes which refer to an entity is one of the challenges introduced by the task. Besides, the BTC collection appears to be noisy and incomplete. For instance, it contains far less Wikipedia entities than those which are the part of the ClueWeb B collection. This may be representative of the situation where entity classes are not that well covered by specialized entity repositories (as opposed to the coverage of the most popular entity classes in Wikipedia).

3.3 Topics and assessments

In order to help participants of 2009 use their previous approaches in the new setup, we use a subset of the 20 topics developed in the 2009 pilot run of the track. We had to exclude 6 topics from this set (#8, #9, #10, #13, #14, and #18) which had either too many additional entities as answers, or whose answer set from 2009 was complete, so could not be extended (for instance, all members of a band were found by participants of REF task in 2009). For each of the remaining 14 topics, the answer entities identified in the 2009 Entity track serve as the list of examples. Both the input entity and the examples were then manually mapped to LOD by track organizers with the help of a baseline entity search system. Entities might be identified by one or more URIs, but the set of URIs corresponding to a given entity is not necessarily complete. Additionally, the target type was mapped to the single most specific class within the DBPedia ontology[4]. An example topic is shown below:

[3] `http://vmlion25.deri.ie/`

[4] `http://wiki.dbpedia.org/Ontology`

```
<query>
   <num>4</num>
   <entity_name>Philadelphia, PA</entity_name>
   <entity_URL>clueweb09-en0011-13-07330</entity_URL>
   <entity_URIs>
      <URI>http://dbpedia.org/resource/Philadelphia</URI>
      <URI>http://sws.geonames.org/4560349/</URI>
   </entity_URIs>
   <target_entity>organization</target_entity>
   <target_type_dbpedia>dbpedia-owl:SportsTeam</target_type_dbpedia>
   <narrative>Professional sports teams in Philadelphia.</narrative>
   <examples>
      <entity>
         <URI>http://dbpedia.org/resource/Philadelphia_Wings</URI>
      </entity>
      <entity>
         <URI>http://dbpedia.org/resource/Philadelphia_KiXX</URI>
         <URI>http://rdf.freebase.com/ns/guid.9202a8c0400064[...]</URI>
      </entity>
      [...]
   </examples>
</query>
```

Relevance judgements were also performed by the track organizers. All submitted runs were assessed up to rank 100 using a binary system of judgments for URIs; names were not evaluated.

One topic had proven too problematic because of the huge set of potentially correct answers (#1). Five more topics had to be excluded because no relevant entities were found for them in the BTC corpus (#2, #3, #6, #16, and #19). This left us with 8 topics in total, listed in Table 3. Similarly to REF, relevant entities were assigned to equivalence classes.

Table 3: ELC topics. #ex is the number of example entities provided and #rel is the number of additional relevant entities identified.

ID	Narrative	#ex	#rel
#4	Professional sports teams in Philadelphia.	6	5
#5	Products of Medimmune, Inc.	2	1
#7	Airlines that currently use Boeing 747 planes.	23	25
#11	Donors to the Home Depot Foundation.	6	8
#12	Airlines that Air Canada has code share flights with.	13	17
#15	Universities that are members of the SEC conference for football.	10	3
#17	Chefs with a show on the Food Network.	22	21
# 20	Scotch whisky distilleries on the island of Islay.	7	1

3.3.1 Qrels

In the qrels file, the fields are:

```
topic doc rel class
```

Table 4: Runs submitted to the ELC task, ordered by MAP. The columns of the table (from left to right) are: runID, group, type of the run (Automatic/Manual), whether the ClueWeb09 collection was used (Yes/No), whether any external resources were used (Yes/No), MAP, R-precision, and number of relevant retrieved results. Highest scores for each metric are in boldface.

Run	Group	Type	CW	Ext.	MAP	rPrec	#rel
KMR5PU	Purdue_IR	A	N	N	**0.2613**	**0.3116**	33
ilpsSetOLnar	UAms	A	N	N	0.1152	0.0899	**43**
ilpsSetOL	UAms	A	N	N	0.1105	0.0947	40
LiraSealClwb	CMU_LIRA	A	Y	N	0.0755	0.0494	15
LiraSealgoog	CMU_LIRA	A	N	Y	0.0228	0.0274	15

Where `topic` denotes the topic ID (corresponds to the `num` field of the topic definition), `doc` is a BTC URI, `rel` is the judgment for the document (0 or 1), and `class` is a class number for the document.

3.3.2 Evaluation metrics

The main evaluation measure we use is Mean Average Precision. We also report on R-Precision (precision at rank R). Relevant entities previously seen in the ranking are considered irrelevant.

Note that evaluation results are not computed using the standard `trec_eval` tool, but a script developed specifically for the ELC task[5].

3.4 Runs and Results

For the ELC pilot task, three groups submitted a total of 5 runs, all of which were automatic. The results are shown in Table 4.

3.5 Approaches

Below are the summaries of approaches, contributed by the participating teams (edited slightly for better presentation).

CMU_Lira The team from CMU (CMU_Lira) focused on Entity List Completion using Set Expansion techniques. Set expansion refers to expanding a partial set of "seed" objects into a more complete set. They propose a two stage retrieval process. The first stage takes the given query_entity and target_entity examples as seeds and does set expansion. In the second stage, candidates generated by first stage are type checked and ranked. The first stage of this approach focuses on recall while the second stage tries to improve precision of the intermediate result list. They have submitted two runs, by doing set expansion on the Web and on the Clueweb corpus. (Dalvi et al., 2011)

Purdue_IR In the entity list completion (ELC) task, we leverage IR techniques to store the semantic data about entities and to retrieve the entities by Indri structured query retrieval language. Furthermore, we perform type matching between the target entity type and the candidate entity type. (Fang et al., 2011)

[5] `http://trec.nist.gov/data/entity/10/eval-entity-elc.pl`

UAms To address ELC we look for entities similar to the given example entities. We find items that are linked to by example entities and consider other entities that link to those items to be candidate entities. For each entity we consider its links as well as the items to which it links. The combination of a link and a linked item forms a link-item pair. Each entity has its set of associated link-item pairs. We rank entities by set overlap between their link-item pairs and the example entity link-item pairs. We then re-rank these intermediate results based on word overlap between the topic narrative and entity link-item pairs. (Bron et al., 2011)

4 Summary

The second edition of the Entity track featured the Related Entity Finding (REF) as the main task: given an input entity, the type of the target entity (person, organization, product, or location), and the relation, described in free text, systems had to return homepages of related entities, and, optionally, the name of the entity.

For the second year of the track, 50 topics were created and assessed. In addition, participants were also requested to generate results for the 20 REF topics from 2009. We had slightly more submissions compared to the previous year (14 vs. 13 participants, 48 vs. 41 runs). This serves as a good motivation to run the task again next year. However, it becomes especially interesting if there are other applications within the same domain which have the potential to attract as many researchers as the REF task.

Entity 2010 also featured a pilot task: Entity List Completion (ELC). ELC is motivated by the same user scenario as REF, but entities are represented by their URIs in a Semantic Web crawl (the Billion Triple Collection), and a small number of example entities are made available as part of the topic definition. Our pilot run of the ELC task was not as popular as REF, probably due to the fact that participation needed a significant additional effort, because of the different nature of the dataset. We plan to run the task again in 2011, so that participants could have enough time to build their systems and process the data.

References

N. Agarwal, H. Bisgin, H. Joshi, M. V. Swamy, S. Wallace, and X. Xu. UALR at 2010 TREC Conference. In *Proceedings of the Nineteenth Text REtrieval Conference (TREC 2010)*, 2011.

K. Balog, A. P. de Vries, P. Serdyukov, P. Thomas, and T. Westerveld. Overview of the TREC 2009 Entity Track. In *Proceedings of the Eighteenth Text REtrieval Conference (TREC 2009)*, 2010.

L. Bonnefoy, P. Bellot, and M. Benoit. LIA-iSmart at TREC 2010 : A Web-oriented Language Modeling Approach for Question Related Entity Finding. In *Proceedings of the Nineteenth Text REtrieval Conference (TREC 2010)*, 2011.

M. Bron, J. He, K. Hofmann, E. Meij, M. de Rijke, M. Tsagkias, and W. Weerkamp. The University of Amsterdam at TREC 2010: Session, Entity and Relevance Feedback. In *Proceedings of the Nineteenth Text REtrieval Conference (TREC 2010)*, 2011.

L. Cao, L. Bai, X. Cheng, J. Guo, H. Xu, Y. Liu, and X. Yu. ICTNET at Entity Track TREC 2010. In *Proceedings of the Nineteenth Text REtrieval Conference (TREC 2010)*, 2011.

B. Dalvi, J. Callan, and W. Cohen. Entity List Completion Using Set Expansion Techniques. In *Proceedings of the Nineteenth Text REtrieval Conference (TREC 2010)*, 2011.

Y. Fang, L. Si, N. Somasundaram, S. Al-Ansari, Z. Yu, and Y. Xian. Purdue at TREC 2010 Entity Track: a Probabilistic Framework for Matching Types between Candidate and Target Entities. In *Proceedings of the Nineteenth Text REtrieval Conference (TREC 2010)*, 2011.

A. Hold, M. Leben, B. Emde, C. Thiele, F. Naumann, W. Barczynski, and F. Brauer. ECIR - A Lightweight Approach for Entity-Centric Information Retrieval. In *Proceedings of the Nineteenth Text REtrieval Conference (TREC 2010)*, 2011.

J. Kamps, R. Kaptein, and M. Koolen. Using Anchor Text, Spam Filtering and Wikipedia for Web Search and Entity Ranking. In *Proceedings of the Nineteenth Text REtrieval Conference (TREC 2010)*, 2011.

Q. Li and D. He. Searching for Entities When Retrieval Meets Extraction. In *Proceedings of the Nineteenth Text REtrieval Conference (TREC 2010)*, 2011.

K. A. Shaik, B. Gosukonda, V. Pande, and V. Varma. IIIT Hyderabad at Entity Track TREC 2010. In *Proceedings of the Nineteenth Text REtrieval Conference (TREC 2010)*, 2011.

O. Vechtomova. Related Entity Finding: University of Waterloo at TREC 2010 Entity Track. In *Proceedings of the Nineteenth Text REtrieval Conference (TREC 2010)*, 2011.

D. Wang, Q. Wu, H. Chen, and J. Niu. A Multiple-Stage Framework for Related Entity Finding: FDWIM at TREC 2010 Entity Track. In *Proceedings of the Nineteenth Text REtrieval Conference (TREC 2010)*, 2011a.

Z. Wang, C. Tang, X. Sun, H. Ouyang, R. Lan, W. Xu, G. Chen, and J. Guo. PRIS at TREC 2010: Related Entity Finding Task of Entity Track. In *Proceedings of the Nineteenth Text REtrieval Conference (TREC 2010)*, 2011b.

Y. Wu, C. Hori, and H. Kawai. NiCT at TREC 2010: Related Entity Finding. In *Proceedings of the Nineteenth Text REtrieval Conference (TREC 2010)*, 2011.

Q. Yang, P. Jiang, C. Zhang, and Z. Niu. Reconstruct Logical Hierarchical Sitemap for Related Entity Finding. In *Proceedings of the Nineteenth Text REtrieval Conference (TREC 2010)*, 2011.

Table 5: All REF runs ordered by NDCG@R. Highest scores for each metric are in boldface.

Run	Group	Type	WP	Ext.	NDCG@R	MAP	rPrec	#rel	#pri
bitRFRun	BIT	M	N	Y	**0.3897**	**0.2876**	**0.3209**	153	319
bitDSHPRun	BIT	A	N	N	0.3694	0.2726	0.3075	150	314
bitDSRRun	BIT	M	N	Y	0.3694	0.2726	0.3075	150	314
FduWimET4	FDWIM2010	A	N	Y	0.3420	0.2223	0.2837	140	**333**
FduWimET2	FDWIM2010	A	Y	Y	0.3382	0.2272	0.2917	120	303
FduWimET3	FDWIM2010	M	Y	Y	0.3376	0.2218	0.2886	116	297
FduWimET1	FDWIM2010	A	Y	Y	0.3259	0.2235	0.2823	83	276
KMR3PU	Purdue_IR	M	Y	Y	0.2917	0.1916	0.2505	93	296
EntityHP1	PITTSIS	M	N	Y	0.2884	0.1664	0.2258	140	278
PRIS2	PRIS	M	N	Y	0.2846	0.1607	0.2296	128	312
EntityHP	PITTSIS	M	Y	Y	0.2837	0.1556	0.2009	**168**	312
KMR1PU	Purdue_IR	A	Y	Y	0.2485	0.1555	0.2099	91	246
PRIS3	PRIS	M	N	Y	0.2160	0.1141	0.1498	141	301
PRIS1	PRIS	M	N	Y	0.2158	0.1180	0.1639	130	310
PRIS4	PRIS	M	N	Y	0.1761	0.0984	0.1361	130	291
SuppHome	NiCT	A	N	Y	0.1696	0.0953	0.1453	74	187
SuppHomeIsA	NiCT	A	N	Y	0.1655	0.0971	0.1446	61	174
ICTNETRun1	ICTNET	A	N	Y	0.1611	0.0839	0.1305	95	173
SIELRUN1	SIEL_IIITH	M	Y	Y	0.1576	0.1019	0.1414	38	198
SIELRUN2	SIEL_IIITH	M	Y	Y	0.1576	0.1019	0.1414	38	198
SIEL10RUN1	SIEL_IIITH	M	Y	Y	0.1576	0.1019	0.1414	38	198
UWAT2	UWaterlooEng	A	N	Y	0.1393	0.0722	0.1223	96	154
UWAT1	UWaterlooEng	A	N	Y	0.1264	0.0608	0.1033	95	151
UWEntTI	UWaterlooEng	A	Y	Y	0.1259	0.0603	0.0974	95	148
SuppIsA	NiCT	A	N	Y	0.1245	0.0703	0.0991	76	143
Supp	NiCT	A	N	Y	0.1237	0.0647	0.0909	85	150
LearnDPI	LIA_UAPV	A	N	Y	0.0766	0.0305	0.0591	72	81
G16	HPI	A	N	Y	0.0745	0.0357	0.0539	27	71
Comp	LIA_UAPV	A	N	Y	0.0737	0.0261	0.0463	74	74
ValueDoc	PITTSIS	M	Y	Y	0.0723	0.0251	0.0500	50	54
ilpsM50agfil	UAms	M	N	Y	0.0718	0.0331	0.0496	36	99
UAcatslinkA	UAmsterdam	M	Y	N	0.0708	0.0485	0.0678	29	84
ilpsM50	UAms	M	N	Y	0.0692	0.0298	0.0455	35	94
UAcatscombB	UAmsterdam	M	Y	N	0.0685	0.0323	0.0452	47	82
G64	HPI	A	N	Y	0.0625	0.0252	0.0500	29	76
RanksDivComp	LIA_UAPV	A	N	Y	0.0610	0.0200	0.0373	76	76
ilpsM50var	UAms	M	N	Y	0.0571	0.0234	0.0375	40	77
UAbaselinkA	UAmsterdam	A	Y	N	0.0496	0.0185	0.0349	34	81
ilpsA500	UAms	A	N	Y	0.0460	0.0178	0.0325	35	88
Div	LIA_UAPV	A	N	Y	0.0428	0.0129	0.0189	76	77
YahooEnHP	PITTSIS	A	N	Y	0.0375	0.0118	0.0229	37	42
UAbaseanchB	UAmsterdam	A	Y	N	0.0314	0.0063	0.0167	47	42
Y64	HPI	A	N	Y	0.0222	0.0055	0.0223	14	39
CARDENSMBLE	CARD_UALR	A	N	N	0.0084	0.0000	0.0003	20	1
CARDSGFCS	CARD_UALR	A	N	N	0.0081	0.0001	0.0006	19	2
CARDFPR	CARD_UALR	A	N	N	0.0077	0.0005	0.0018	18	3
CARDHITS	CARD_UALR	A	N	N	0.0070	0.0001	0.0003	24	2
B64	HPI	A	N	Y	0.0178	0.0044	0.0122	12	30
Median					0.1252	0.0628	0.0983	74	149