

CLEF 2005
Overview of Multilingual Retrieval and
Merging Track

Gareth J. F. Jones

Dublin City University

Overview

- Track Definition
- Test Collection
- Participants
- Results
- Conclusions
- Outlook

Track Definition

Two subtasks offered at CLEF 2005:

- Multi-8 Two Years On: full multilingual retrieval task.
 - repeat CLEF 2003 Multi-8 task to assess progress in multilingual information retrieval.
- Multi-8 Merging Only: participants merge provided bilingual ranked lists.
 - aim to encourage detailed exploration of merging strategies without needing to perform retrieval stage.

Test Collection

- Multi-8 target document collections in the following 8 languages: Dutch, English, Finnish, French, German, Italian, Spanish, and Swedish.
 - appox. 1.5 million documents.
- Original 60 topics split as follows:
 - topics 141-160 training topic set.
 - topics 161-200 test topic set.
- No new relevance assessments carried out.

Test Collection

- Standard ranked retrieval lists from:
 - University of Neuchatel
 - Hummingbird

Bilingual runs using English topics.

- Enables comparison of merging strategies across multiple retrieval systems.

Participants

Multi-8 Two Years On

CMU - Lang.Tech. (USA)*
Daedalus & Madrid Univs (Spain)
U.Geneva - Inf.Systems (Switzerland)*
U.Jaen - Intell.Systems (Spain)

Multi-8 Merging Only

CMU - Lang.Tech. (USA)*
Dublin City U. - Computing (Ireland)
U.Jaen - Intell.Systems (Spain)

Multi-8 Results

	Task	Top Perf.	Av. Prec.
CLEF 2005	2 Years On	CMU	0.449
	Merging Only	CMU	0.412
CLEF 2003	Multi-8	Berkeley	0.388

- Results for submitted CLEF 2003 Multi-8 runs recalculated for topics 161-200.
- Absolute values change slightly, but ranking of runs is preserved.

Conclusions

- More sophisticated merging strategies can improve on results obtained with earlier retrieval runs.
- Use of current retrieval systems can offer further offer further improvements on previous retrieval performance.

But:

- group ranked first at CLEF 2005 didn't participate in CLEF 2003; maybe their retrieval system was already better.
- CLEF 2005 had a training topic set; this was not available for original CLEF 2003 participants.

Outlook

- Improvements in merging can be achieved over previous results.
 - how much scope is there for further improvements without overfitting?
 - might extended format ranked lists offer further improvement in merging performance?
- Performance from earlier CLEF workshops can (at least in some cases) be improved.
 - should we work on improving monolingual, bilingual and multilingual retrieval performance using existing test collections?

Acknowledgements

Thanks for the provision of ranked lists:

- Stephen Tomlinson
- Jacques Savoy