

Group Creation and Complex Queries

The interface allows to create groups of systems, documents and/or metrics. The purpose of this feature is to facilitate the comparison between types of systems (e.g., statistical vs. rule-based) or metrics (e.g., lexical vs. syntactic) or even, groups of documents that belong to different domains.

Let consider two metric groups: $LEX = \{BLEU, NIST\}$ and $SYN = \{SP - Op, DP - Or\}$. The following query

```
(UPC : SYN > AVG AND UPC : LEX < AVG)
OR (UPC : LEX > AVG AND UPC : SYN < AVG)
```

returns all segments belonging to the translation system UPC that have good scores for lexical metrics and bad scores for syntactic ones, or viceversa. Note that the structure defined at the previous point for systems, metrics and documents, is the same at group-level, e.g., $news : LEX : RBMT > AVG$, where RBMT is a group of rule-based systems.

On-line Interface and Export

*t*SEARCH lets the user to navigate the results of the search over the automatic translations and their evaluations. Three different views organize the segments according to the user preferences:

All segments this view shows all segments and the metric scores involved in the query.

Grouped by system groups the segments by system name and, for each system, by document name.

Grouped by segment this view offers the segment organization, which facilitates the comparison between several translations, the reference and the source

Finally, the search output can be exported as an XML file. Each view offers the chance to do it at different levels. Users can even get all segments without being grouped at the main view.

*t*Search

Flexible and Fast Search over
Automatic Translations for
Improved Quality/Error Analysis

Meritxell Gonzàlez, Laura Mascarell
and Lluís Màrquez

TALP Research Center
Universitat Politècnica de Catalunya

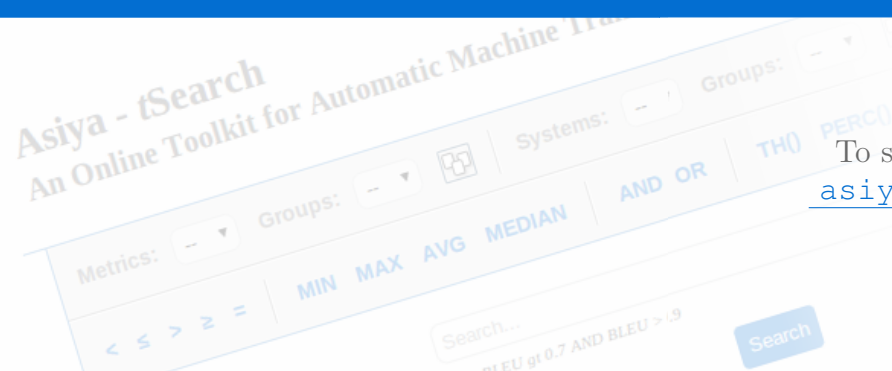
{mgonzalez, lmascarell, lluis}@lsi.upc.edu



Helping MT developers

Analyze efficiently

Scan me!



To start using *t*SEARCH visit
asiya.lsi.upc.edu/demo/



Why *t*Search?

*A powerful tool for
fast and flexible
MT error analysis*

*t*SEARCH is an on-line application that provides mechanisms for doing complex searches over a collection of translation cases evaluated with a large set of diverse measures. *t*SEARCH aids the error analysis stage of machine translation development, it helps to identify and analyze system's weaknesses.

The flexible query language allows to find translation examples matching a combination of numerical and structural features associated to translation quality measures.

*t*SEARCH has been built on top of ASIYA, an open-source toolkit for MT evaluation. *t*SEARCH uses the evaluation results from ASIYA and, along with the ASIYA on-line GUI, it makes possible a graphical visualization and interactive access to the evaluation results.

Participate!

We are conducting an evaluation of the user interface and we need your help!

The goal of the test is to assess the usability of the on-line interface of Asiya and the *t*Search tool. It consists of three easy steps: first, a simple training; second, the resolution of the test scenario; and third, the completion of a short questionnaire.

Please, participate in the experiment at:
asiya.lsi.upc.edu/demo/HCItest/

Metric-based Queries

Queries are divided between those that use comparison operators and those that involve a range of values:

Arithmetic comparison	BLEU > 0.4
	BLEU > TH(40)
	BLEU le MEDIAN
Range of values	BLEU IN [0.2, 0.3)
	BLEU IN (TH(20), TH(40)]
	BLEU IN Q(4)
	BLEU IN PERC(2, 10)

Arithmetic comparison queries let obtain all segments scored above/below a value for a concrete metric. Such value can be a real number or also a statistical variable such as minimum MIN, maximum MAX, median MEDIAN, average AVG or the threshold function TH().

We have implemented statistical functions such as the quartile function Q() or the percentile PERC(n,M), which returns all the segments with a score in the n^{th} part, when the range of scores is divided in M parts of equal size.

The queries in the table above are applied at segment level. However, applying them at system and document-level is as easy as specifying the system and/or document names: the query UPC : BLEU > AVG returns the segments concerning the UPC translations with a BLEU score above its average.

Linguistic Elements-based Queries

Linguistic-wise queries allow to find the segments that match some criteria related to the linguistic features computed by ASIYA:

N-grams	LE [SP (NP, *, PP)]
	LE [CP (VP, PP)]
	LE [lemma (be), CP (VP, PP)]
	LE [pos (DT, JJ, *)]
Semantic Roles	LE [SR (ask, A1, AM-TMP)]
	LE [SR (*, A1, AM-TMP)]
	LE [SR (**, A1, AM-TMP)
Dependency relationships	LE [DP (N, nsubj, V)
	LE [DP (N, nsubj, V, dep, V)]
	LE [DP (*, nsubj, *)

We have implemented queries that match N-grams of lemmas `lemma`, parts-of-speech `pos` and items of shallow SP or constituent parsing CP, dependency relations DP and semantic roles SR. The DP function allows specifying a structure composition criterion (i.e., the categories of two words and their dependency relationship) and even a chain of relations. The SR function obtains the segments that match a verb and its list of arguments.

The use of the asterisk symbol substitutes any value, e.g., LE[SP(NP, *, PP), DP(*, *, V)].

As metric-based queries these can be also applied at system/document level, e.g., UPC : LE[SP(NP, PP)].