This work presents an online graphical interface to access ASIYA, an open source toolkit to evaluate automatic translations using an heterogeneous set of metrics and meta-metrics.

1. To allow MT developers to evaluate their test beds using a large set of metric scores
2. To detect and analyze the errors of the MT systems using just their Internet browsers
3. To help developers to understand the strengths and weaknesses of the evaluation measures

The Online Graphical Interface

1. The online form allows to upload files and select the required options.
2. Asiya generates a number of metric-dependent information...
3. ...that produce an interactive evaluation report.
4. The visualization of the linguistic information is useful for MT developers, such as interactive annotations and parse trees:

The ASIYA Toolkit

- More than 500 metric variants for MT evaluation
- Various similarity principles: precision, recall and overlap
- Different linguistic layers:
  - Lexical similarity: based on n-gram similarity and edit distance based on word form, e.g., PER, TER, WER, BLEU, NIST, GTM, METEOR
  - Syntactic similarity: based on part-of-speech tags, base phrase chunks, dependency and constituency trees e.g., SP-Overlap-POS, DP-HWCM, CP-STTM
  - Semantic similarity: based on named entities, semantic roles and discourse representation e.g., NE-Overlap, SR-Overlap, DRS-Overlap

Future Work

- Improve usability of the interface, e.g., allow input texts, dote the parse trees with more interactions.
- Create a database to save test sets and results.
- Show word alignments and use them to calculate metrics under a new principle.
- Detect and classify errors automatically.
- Create a search engine to filter results and obtain specific good/bad examples from the test sets.

This research has been partially funded by the Spanish Ministry of Education and Science [OpenMT-2, TIN2009-14875-C02] and the European Community's Seventh Framework Programme under grant agreement numbers 247762 (PAUST project), FP7-ICT-2009-4-247762 and 247914 (MOLTO project), FP7-ICT-2009-4-247914.