1 MEADeval

MEADeval is a toolkit for evluating MEAD- and DUC-style extracts. It supports a number of standard metrics, as detailed below. MEADeval is a collection of Perl modules, with example scripts that detail how to use the modules. Users may want to modify the scripts to fit the scripts of modules to fit their needs.

This document assumes familiarity with the MEAD documentation, available at http://www.clsp.jhu.edu/ws2001/groups/asmd/. Also, many references are made to the Document Understanding Conference (DUC) and some supporting document formats and documentation. The web site for DUC2002 is http://www-nlpir.nist.gov/projects/duc/. The DUC web site contains information about the data formats and has some examples.

1.1 Supported Evaluation Metrics

The metrics for extracts only are:

- precision
- recall
- precisionw

  Like precision, but normalized by the length (in words) of each sentence.
- recallw

  Like recall, but normalized by the length (in words) of each sentence.
- kappa
- relative utility

  A paper on relative-utility is available at http://perun.si.umich.edu/clair/meadeval/ru.ps and is included in the www directory of the distribution.
- normalized relative utility

The metrics for general text are:

- unigram overlap
- bigram overlap
- cosine
- simple cosine

  Cosine without adjustments for IDF.

1.2 Downloading and Installation

Download, unzip, and untar meadeval.tar.gz. This results in a directory named meadeval. Within meadeval are:

- /bin

  contains test scripts
- /data

  contains sample extracts and other data files.

  mead-cluster125.extract is a MEAD-style extract for cluster numbered 125. cluster125.sentjudge is a sentjudge file for the same cluster. duc1-cluster1.extract and duc2-cluster1.extract are two DUC-style extracts for cluster number 1.
- /lib

  contains the perl modules for evaluation metrics

- /etc

  contins perl DBM files for use by content-based metrics.

Let $MEADEVAL be the full path to the meadeval directory that was created in the untarring/zipping process (including the meadeval directory). Before using meadeval, perl must know where to look for modules, so, add the meadeval/lib directory to the PERL5LIB environment variable. For example:

export PERL5LIB=$PERL5LIB:$MEADEVAL/lib

1.3 Supported Data Formats

MEADeval operates mainly on extract files. An extract file describes the sentences contained in an extractive summary: which document each came from and the number of each sentence within the source document. Both MEAD- and DUC-style extracts are accepted by MEADeval.

Additionally, the Relative Utility metric requires relative utility judgements by one or more human judges. These are held in a SentJudge file. This file format is described in the MEAD documentation.

1.4 MEADeval command-line scripts

Command-line scripts are in the $MEADEVAL/bin directory. Remember to add the $MEADEVAL/lib directory to the PERL5LIB environment variable before attempting to run the scripts.

NOTE: These scripts are meant as examples only and Individual users will likely want to modify one or both of these scripts to suit their needs.

1.4.1 meadeval.pl and duc-meadeval.pl

These two scripts compute and display most of the evaluation metrics MEADeval offers. The difference is that meadeval.pl works on MEAD-style extracts and duc-meadeval.pl requires DUC-style extracts.

There is only one example MEAD-style extract included in the documenation, so meadeval.pl is mostly a placeholder, that is used exactly like duc-meadeval.pl, but with MEAD-style extracts.

NOTE: since MEAD-style extracts do not contain the text of the extract, content-based measures will not work on these extracts.

To run duc-meadeval.pl:

A number of statistics are printed to the screen.

1.4.2 relative-utility.pl

This script computes and displays the relative utility of an extract, based on utility judgements from one or more judges.

Usage:

NOTE: the extracts for this script must be MEAD-style extracts.

1.5 MEADeval API

What follows is a brief decription of the modules included in the MEADeval distribution. For a complete description of the MEADeval API, see http://perun.si.umich.edu/clair/meadeval/

- Essence::IDF

  Contains methods for accessing IDF DBM files. You shouldn't need to use it...

- DUC::Extract

  Represents a DUC-style extract.

- MEAD::Extract

  Represents a MEAD-style extract.

- MEAD::Evaluation

  Contains the code for the evaluation metrics supported by MEADeval.

- MEAD::SentJudge

  Represents a SentJudge file. A SentJudge file contians a rating (1 to 10) by each judge, for each sentence in the cluster, of how important the sentence is to the cluster.

1.6 Credits

MEADeval was written by Adam Winkel and Dragomir Radev.

Thanks to Simone Teufel, Horacio Saggion, Hong Qi, and John Blitzer who all provided code and/or expertise for MEADeval.