Using Resnik Script

Resnik is a biological analysis script. It is a command-line utility that leverages an existing library, fastsemsim, created by Marco Mina.

Location

Resnik is located in the repository /extra/wayne0/preserve/resnik/. It is titled **resnik3.py**. It is also located within the resnik folder of the sana respository.

Before Running the Script

Please unzip gene2go.gz and go.obo.gz within sana/resnik. The files are zipped to preserve space.

Command-line Arguments

Required Arguments:

`-g1` , `--graph1` — 1st graph file, LEDA format.
`-g2` , `--graph2` — 2nd graph file, LEDA format. Second graph must be smaller than the 1st

Another required argument is an alignment file. There are two formats accepted: SANA and a two-column alignment file.

For a two-column alignment file, use `-a` or `--alignment-file`.
For a SANA-outputted alignment file, use `-so` or `--sana-out`

Default Arguments:

`-t`, `--taxon` — List of NCBI taxon ids to include, separated by dashes. If it it not specified, the script will include all taxons.

`--mode` — Choice of mixing semantic scores. Choices are *max*, *avg*, and *BMA*. Default choice is max.

`-ec`, `--evidence-codes` — Evidence codes to include/exclude, separated by dashes. Default is to include all evidence codes.

`-i` — Default is to exclude the evidence codes listed in `-ec`. Choosing this option will *include* the evidence codes.

`-oc`, `--ontology-codes` — Ontology codes to include, separated by dashes. Default is to include all.

`-o`, `--output` — If you want the script to write the output to a file, use this option. If not, the program will print to stdout.

`-v`, `--verbose` — Program will print out verbose output

`--semsim` – Semantic measure to use. Choices can be viewed by running `resnik3.py -h`

Examples

```
resnik3.py -g1 RNorvegicus.gw -g2 SCerevisiae.gw -so
sana.out -t Athaliana-RNorvegicus --mode avg
```

This command uses RNorvegicus.gw and SCerevisiae.gw as the 1st and 2nd graphs. The alignment file is sana.out. The taxons included

are the codes listed in Athaliana.txt and RNorvegicus.txt within sana/resnik/taxons. Between each alignment pair, the average of all semantic scores are calculated.

```
resnik3.py -g1 RNorvegicus.gw -g2 Scerevisiae.gw -a
my_alignment.file -ec IEA -oc BP -o my_output.txt -v --
semsim Jiang-Conrath
```

This command uses the two-column alignment file *my_alignment.file*. Excludes the evidence code IEA. Only tests for ontology code *BP*. Writes output to *my_output.txt*. Uses semantic similarity measure Jiang-Conrath.