NAME

      RDKitEnumerateStereoisomers.py - Enumerate stereoisomers of molecules

SYNOPSIS

      RDKitEnumerateStereoisomers.py [--discardNonPhysical <yes or no>] [--infileParams <Name,Value,...>] [--mode <UnassignedOnly or All>] [--maxIsomers <number>] [--outfileParams <Name,Value,...>] [--overwrite] [-w <dir>] -i <infile> -o <outfile>

      RDKitEnumerateStereoisomers.py -h | --help | -e | --examples

DESCRIPTION

      Perform a combinatorial enumeration of stereoisomers for molecules around all or unassigned chiral atoms and bonds.

      The supported input file formats are: Mol (.mol), SD (.sdf, .sd), SMILES (.smi, .csv, .tsv, .txt)

      The supported output file format are: SD (.sdf, .sd), SMILES (.smi)

OPTIONS

      -d, --discardNonPhysical <yes or no> [default: yes]

            Discard stereoisomers with non-physical structures. Possible values: yes or no. The non-physical nature of a stereoisomer is determined by embedding the structure to generate a conformation for the stereoisomer using standard distance geometry methodology.

            A word to the wise from RDKit documentation: this is computationally expensive and uses a heuristic that could result in loss of stereoisomers.

      -e, --examples

            Print examples.

      -m, --mode <UnassignedOnly or All> [default: UnassignedOnly]

            Enumerate unassigned or all chiral centers. The chiral atoms and bonds with defined stereochemistry are preserved.

      --maxIsomers <number> [default: 50]

            Maximum number of stereoisomers to generate for each molecule. A value of zero indicates generation of all possible steroisomers.

      -h, --help

            Print this help message.

      -i, --infile <infile>

            Input file name.

      --infileParams <Name,Value,...> [default: auto]

            A comma delimited list of parameter name and value pairs for reading molecules from files. The supported parameter names for different file formats, along with their default values, are shown below:

```
SD, MOL: removeHydrogens,yes,sanitize,yes,strictParsing,yes
SMILES: smilesColumn,1,smilesNameColumn,2,smilesDelimiter,space,
    smilesTitleLine,auto,sanitize,yes
```

            Possible values for smilesDelimiter: space, comma or tab.

      -o, --outfile <outfile>

            Output file name.

      --outfileParams <Name,Value,...> [default: auto]

            A comma delimited list of parameter name and value pairs for writing molecules to files. The supported parameter names for different file formats, along with their default values, are shown below:

```
SD: compute2DCoords,auto,kekulize,no
SMILES: kekulize,no,smilesDelimiter,space, smilesIsomeric,yes,
    smilesTitleLine,yes
```

            Default value for compute2DCoords: yes for SMILES input file; no for all other file types.

      --overwrite

Overwrite existing files.

-w, --workingdir <dir>

Location of working directory which defaults to the current directory.

EXAMPLES

To enumerate only unassigned atom and bond chiral centers along with discarding of non-physical structures, keeping a maximum of 50 stereoisomers for each molecule, and write out a SMILES file, type:

```
% RDKitEnumerateStereoisomers.py  -i Sample.smi -o SampleOut.smi
```

To enumerate only unassigned atom and bond chiral centers along with discarding any non-physical structures, keeping a maximum of 250 stereoisomers for a molecule, and write out a SD file, type:

```
% RDKitEnumerateStereoisomers.py  --maxIsomers 0 -i Sample.smi
  --maxIsomers 250 -o SampleOut.sdf
```

To enumerate all possible assigned and unassigned atom and bond chiral centers, without discarding any non-physical structures, keeping a maximum of 500 stereoisomers for a molecule, and write out a SD file, type:

```
% RDKitEnumerateStereoisomers.py  -d no -m all --maxIsomers 500
  -i Sample.smi -o SampleOut.sdf
```

To enumerate only unassigned atom and bond chiral centers along with discarding of non-physical structures, keeping a maximum of 50 stereoisomers for each molecule in a CSV SMILES file, SMILES strings in column 1, name in column 2, and write out a SD file with kekulization, type:

```
% RDKitEnumerateStereoisomers.py  --infileParams
  "smilesDelimiter,comma,smilesTitleLine,yes,smilesColumn,1,
  smilesNameColumn,2" --outfileParams "compute2DCoords,yes,
  kekulize,yes" -i SampleSMILES.csv -o SampleOut.sdf
```

AUTHOR

Manish Sud(msud@san.rr.com)

SEE ALSO

RDKitConvertFileFormat.py, RDKitEnumerateCompoundLibrary.py, RDKitGenerateConformers.py, RDKitGenerateMolecularFrameworks.py

COPYRIGHT

Copyright (C) 2019 Manish Sud. All rights reserved.

The functionality available in this script is implemented using RDKit, an open source toolkit for cheminformatics developed by Greg Landrum.

This file is part of MayaChemTools.

MayaChemTools is free software; you can redistribute it and/or modify it under the terms of the GNU Lesser General Public License as published by the Free Software Foundation; either version 3 of the License, or (at your option) any later version.