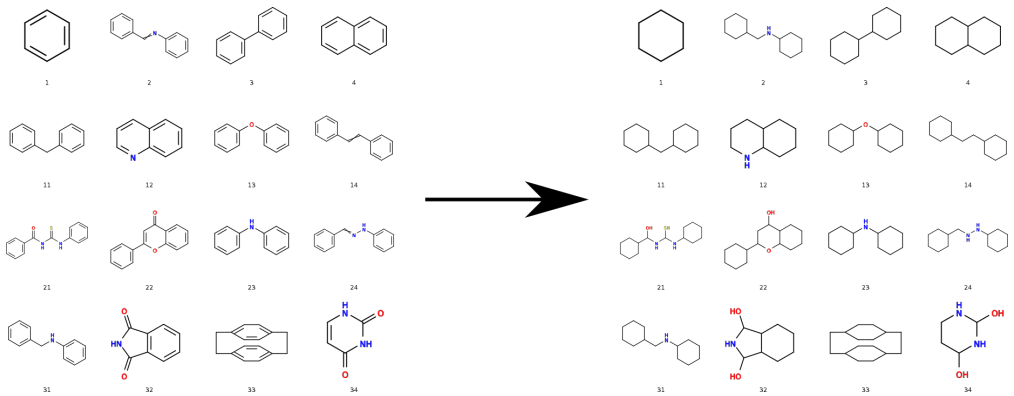
1 Background

The Bemis-Murcko scaffold[1] provided by `DataWarrior`[2] retains information about bond order and chirality. Sometimes, however, it suffices to retain only atom connectivity, like an assumption «there are only single bonds». Note `DataWarrior` equally offers the export of Bemis-Murcko skeleton, however this simplifies e.g. the scaffold about an imidazole into one of cyclopentane.



2 Typical use

The script runs in the CLI of Python. File `listing_file.txt` contains the SMILES to work with:

```
python saturate_MurckoScaffolds.py [listing_file.txt]
```

This generates `saturated_Murcko_scaffold_sat.smi` as permanent record, which shares the file name with the input file *with the addition of* `_sat` *(like saturated).* Regardless of the three-character file extension of the input, the plain ASCII-list of SMILES is recorded in a file of extension `.smi`, accepted recognized e.g., by `openbabel` to contain this specific information.

While intended to work with the current branch of Python 3.6+, the script equally works well with the legacy branch of Python 2, e.g., Python 2.7.16.

3 Example

For a collection of organic materials, the Bemis-Murcko scaffolds was extracted with `DataWarrior` (then release 5.0.0 for Linux, January 2019) as listing `test_input.smi.txt` including higher bond orders (see folder `test_data`). The «artificial saturation» was triggered by

```
1 python saturate_MurckoScaffolds.py test_input.smi
```

to yield `test_input_sat.smi`. The effect of this operation is easy to recognize while comparing the scaffold lists (fig. 1) in a difference view of the two `.smi` files.

[1] Bemis GW, Murcko MA *J. Med. Chem.* 1996, **39**, 2887-2893, doi 10.1021/jm9602928

[2] Sander T, Freyss J, von Korff M, Rufener C, *J. Chem. Inf. Model.* 2015, **55**, 460-473, doi 10.1021/ci500588j. The program, (c) 2002–2019 by Idorsia Pharmaceuticals Ltd., is freely available under `http://www.openmolecules.org`. For the source code, see https://github.com/thsa/datawarrior.

Figure 1: Difference view of the SMILES strings of a Murcko scaffold *prior* (left hand column) and *after* an «artificial saturation» (right hand column). Note the removal of explicit bond order indicators, e.g. double bond (equality sign), triple bond bond (octohorpe), or about implicit aromatization (lower case → upper case for atoms of carbon, nitrogen, oxygen (depicted); or sulfur (not depicted). At the same time, stereochemical indicators are removed, too (e.g., slashes in lines #018 and #019, @-signs line #025).

Subsequently, `openbabel`[3] was used to illustrate the work performed. While eventually automated (cf. script `test_series.py`, deposit in folder `test_data`), instructions delivered to `openbabel` on the command line follow the pattern of

```
obabel -ismi test_input.smi -O test_input_color.svg -xc10 -xr12 -xl --addinindex
```

to generate a `.svg` file (vector representation), or

```
obabel -ismi test_input_sat.smi -O test_input_sat_color.png -xc10 -xr12 -xl
↪  --addinindex -xp 3000
```

to generate a bitmap `.png`. The instructions share the definition of data input (`-ismi`) – eventually *a list of SMILES*, rather than only one SMILES string – and the definition of the output format based on the extension of the file to be written (`.svg`, or `.png`). Both put the entries into an array of 10 columns and 12 rows (`-xc`, `xr`) and label them in the consecution of their appearance in the input files (`--addinindex`) as inner index. Contrasting to the vector output, in case of `.png`, an explicit scaling by a factor of 10 (`-xp` parameter) was required to yield an intelligible illustration.

It is remarkable how well `openbabel`'s algorithm displays the molecular structures with advanced motifs (e.g., the scaffold of cyclophane [entry #33 in the synopses, `test_data` folder], sparteine [#38], or adamantane [#50]), some equally depicted in the first illustration of this guide.

4 License

[3] www.openbabel.org. The script initially was developed for and tested with openbabel (release 2.4.1; Nov 12, 2018) and Python 2.7.17 provided by Linux Xubuntu 18.04.2 LTS. It equally works with Python 3.9.1+ (released January 20, 2021) and openbabel (release 3.1.1 by January 6, 2021) as provided in Debian 10.