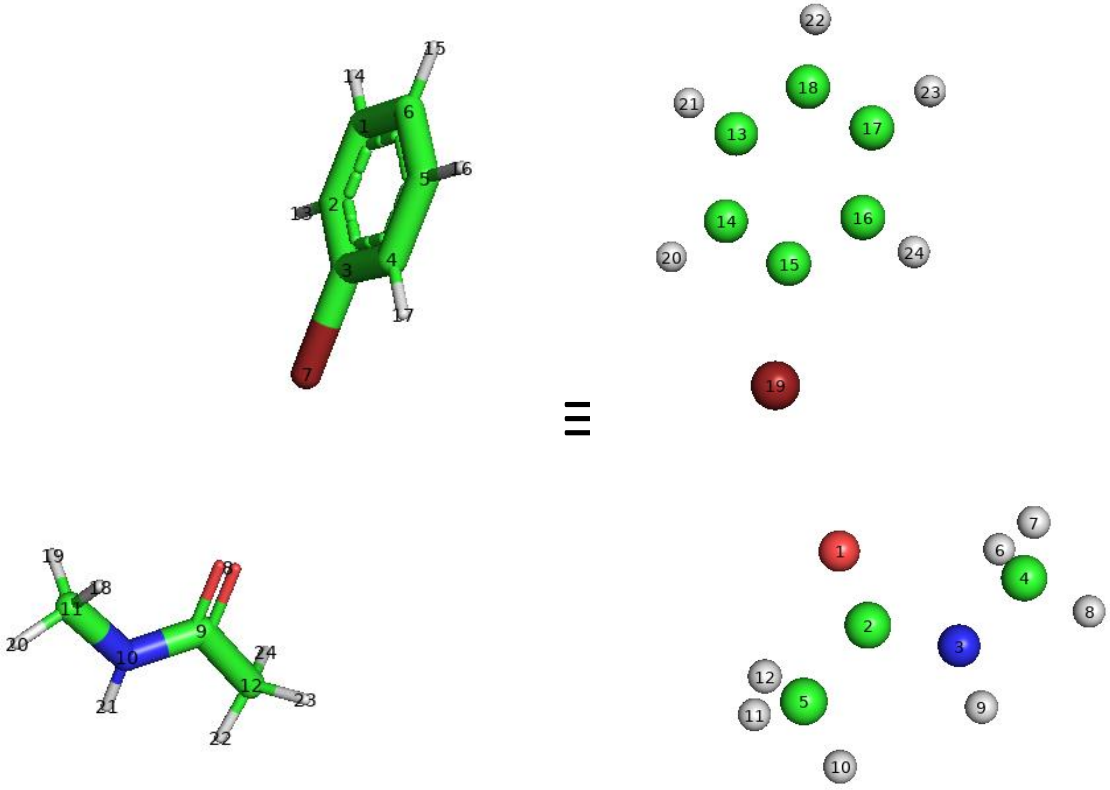
Unique Atom Indexing for arbitrary Molecules

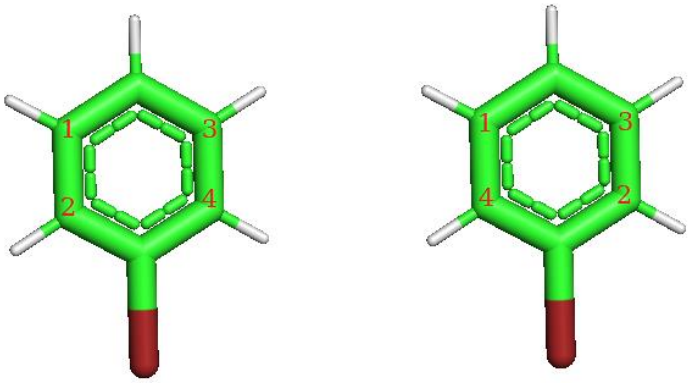
Goal:



We would like to map *chemically identical* atoms to one another, assigning unique indices to every atom. The indexing (or hashing) function should have the following properties:

- In every molecule (file) every index is unique.
- For every atom, if there exists at least one chemically equivalent partner in the other molecule, at least one such partner has the same index.
- Applying the indexing to the same file multiple times yields the same indices.

- Indexing should be consistent with molecular structure. We want to avoid:



Extensibility:

- Application to non-equivalent Molecules
- Runtime improvements (Implementations in C, callable from python?)
- Application to Macromolecules => Structure Alignment

Area:

- Algorithmics
- Graphs
- 3D Structure
- RDKit / PyMOL

Starting Points:

- RDKit:
  - Chem.MolFromPDB, Chem.MolFromXYZ, Chem.SDMolSupplier
  - atom.GetNeighbors
- Floyd-Warshall Algorithm