

chem-bla-ics

Finding differences between IChemObjects #2

Egon Willighagen 

Published June 3, 2008

Citation

Willighagen, E. (2008). Finding differences between IChemObjects #2. In *chem-bla-ics*. chem-bla-ics. <https://doi.org/10.59350/9tk1a-d6c07>

Keywords

Cdk, Qsar

Copyright

Copyright © Egon Willighagen 2008. Distributed under the terms of the [Creative Commons Attribution 4.0 International License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

chem-bla-ics

CDK QSAR descriptors are not allowed to change the input [molecule|atom|bond], and I recently added a unit tests (rev [11138](#)) for that to the abstract class [AtomicDescriptorTest](#).

After some code clean up of the diff module code earlier this morning (in anticipation of the rain stopping), I applied this patch (rev [11269](#)) that `noModification` unit test:

```
public void testCalculate_NoModifications() throws Exception {
    IAtomContainer mol = someoneBringMeSomeWater();
    IAtom atom = mol.getAtom(1);
-   String priorString = atom.toString();
+   IAtom clone = (IAtom)mol.getAtom(1).clone();
    descriptor.calculate(atom, mol);
-   String afterString = atom.toString();
+   String diff = AtomDiff.diff(clone, atom);
    assertEquals(
-       "The descriptor must not change the passed bond in any respect.",
-       priorString, afterString
+       "The descriptor must not change the passed bond in any respect, but found this diff:
+       0, diff.length()
    );
}
```

This is a nice example of where the new [diff module](#) is useful. Instead of dumping to long `IAtom.toString()`s, the output now gives output like:

```
AtomDiff(AtomTypeDiff(, NULL/H, NC:0/1, V:0/1))
```

This indicates (yes, a bit cryptic) that the formal neighbor count (NC) and the valence (V) fields have been modified, in addition to that first field, which I don't know what it refers too. Indeed, the output still needs a bit more tuning :)