Subject: Re: extension of substructure definition
Posted by thomas on Thu, 24 Aug 2023 14:19:33 GMT
View Forum Message <> Reply to Message

Dear Norwid,

The explanation is the following:

DataWarrior distinguishes between molecules and substructures. Molecules are considered complete structures with all open valences meant to be filled with hydrogen atoms. Molecules cannot carry query features on their atoms and bonds. Substructures (also caled fragments), however, have open valences and may have broadening or restricting query features (e.g. atom must be aromatic, or bond may be singe or double). They also may contain exclude groups.

If you open a molecule (even an empty one) in the DataWarrior editor, you cannot add query features, because they are not allowed. When editing a substructure, you may add and change query features. Molecules and substructures are logically different species, need different handling, and are differently visualized.

An empty DataWarrior structure column contains molecules as default even if a cell is empty. If an SD-file is read, then the individual structure cells contains molecules except for those entries, where the molfile read contained a query feature and, thus, implicitly defined the structure to be a substructure. This way a structure column may contain a mix of molecules and substructures. Any new row will contains molecules, unless the column is marked to contains substructures, which usually is not the case.

You may mark a structure column manually to contain substructures by adding the column property isFragment=true. This can be done as macro task or directly in a text editor. The relevant part would then look like this:

<columnName="Structure">
<columnProperty="specialType idcode">
<columnProperty="isFragment true">

Note that the white space between 'isFragment' and 'true' must be a TAB.

Best wishes,

Thomas