NAME

TopologicalPharmacophoreAtomTripletsFingerprints

SYNOPSIS

use Fingerprints::TopologicalPharmacophoreAtomTripletsFingerprints;

use Fingerprints::TopologicalPharmacophoreAtomTripletsFingerprints qw(:all);

DESCRIPTION

TopologicalPharmacophoreAtomTripletsFingerprints [Ref 66, Ref 68-71] class provides the following methods:

new, GenerateFingerprints, , GetDescription, GetAtomTripletIDs, SetAtomTypesToUse, SetDistanceBinSize, SetMaxDistance, SetMinDistance, StringifyTopologicalPharmacophoreAtomTripletsFingerprints

TopologicalPharmacophoreAtomTripletsFingerprints is derived from Fingerprints class which in turn is derived from ObjectProperty base class that provides methods not explicitly defined in TopologicalPharmacophoreAtomTripletsFingerprints, Fingerprints or ObjectProperty classes using Perl's AUTOLOAD functionality. These methods are generated on-the-fly for a specified object property:

```
Set<PropertyName>(<PropertyValue>);
$PropertyValue = Get<PropertyName>();
Delete<PropertyName>();
```

Based on the values specified for AtomTypesToUse, pharmacophore atom types are assigned to all non-hydrogen atoms in a molecule and a distance matrix is generated. Using MinDistance, MaxDistance, and DistanceBinSize values, a binned distance matrix is generated with lower bound on the distance bin as the distance in distance matrix; the lower bound on the distance bin is also used as the distance between atom pairs for generation of atom triplet identifiers.

A pharmacophore atom triplets basis set is generated for all unique atom triplets constituting atom pairs binned distances between --MinDistance and --MaxDistance. The value of --UseTriangleInequality determines whether the triangle inequality test is applied during generation of atom triplets basis set. The lower distance bound, along with specified pharmacophore types, is used during generation of atom triplet IDs.

```
Let:

P = Valid pharmacophore atom type

Px = Pharmacophore atom x
Py = Pharmacophore atom y
Pz = Pharmacophore atom z

Dmin = Minimum distance corresponding to number of bonds between two atoms
Dmax = Maximum distance corresponding to number of bonds between two atoms
D = Distance corresponding to number of bonds between two atom

Bsize  = Distance bin size
Nbins = Number of distance bins

Dxy = Distance or lower bound of binned distance between Px and Py
Dxz = Distance or lower bound of binned distance between Px and Pz
Dyz = Distance or lower bound of binned distance between Py and Pz

Then:

PxDyz-PyDxz-PzDxy = Pharmacophore atom triplet IDs for atom types Px,
                    Py, and Pz

For example: H1-H1-H1, H2-HBA-H2 and so on.

For default values of Dmin = 1 , Dmax = 10 and Bsize = 2, the number of
distance bins, Nbins = 5, are:

[1, 2] [3, 4] [5, 6] [7, 8] [9 10]

and atom triplet basis set size is 2692.

Atom triplet basis set size for various values of Dmin, Dmax and Bsize in
conjunction with usage of triangle inequality is:
```

Dmin	Dmax	Bsize	UseTriangleInequality	TripletBasisSetSize
1	10	2	No	4960
1	10	2	Yes	2692 [Default]
2	12	2	No	8436
2	12	2	Yes	4494

Using binned distance matrix and pharmacohore atom types, occurrence of unique pharmacohore atom triplets is counted.

The final pharmacophore atom triples count along with atom pair identifiers involving all non-hydrogen atoms constitute pharmacophore topological atom triplets fingerprints of the molecule.

For *ArbitrarySize* value of AtomTripletsSetSizeToUse, the fingerprint vector correspond to only those topological pharmacophore atom triplets which are present and have non-zero count. However, for *FixedSize* value of AtomTripletsSetSizeToUse, the fingerprint vector contains all possible valid topological pharmacophore atom triplets with both zero and non-zero count values.

The current release of MayaChemTools generates the following types of topological pharmacophore atom triplets fingerprints vector strings:

```
FingerprintsVector;TopologicalPharmacophoreAtomTriplets:ArbitrarySize:
MinDistance1:MaxDistance10;696;NumericalValues;IDsAndValuesString;Ar1-
Ar1-Ar1 Ar1-Ar1-H1 Ar1-Ar1-HBA1 Ar1-Ar1-HBD1 Ar1-H1-H1 Ar1-H1-HBA1 Ar1
-H1-HBD1 Ar1-HBA1-HBD1 H1-H1-H1 H1-H1-HBA1 H1-H1-HBD1 H1-HBA1-HBA1 H1-
HBA1-HBD1 H1-HBA1-NI1 H1-HBD1-NI1 HBA1-HBA1-NI1 HBA1-HBD1-NI1 Ar1-...;
46 106 8 3 83 11 4 1 21 5 3 1 2 2 1 1 1 100 101 18 11 145 132 26 14 23
28 3 3 5 4 61 45 10 4 16 20 7 5 1 3 4 5 3 1 1 1 1 5 4 2 1 2 2 2 1 1 1
119 123 24 15 185 202 41 25 22 17 3 5 85 95 18 11 23 17 3 1 1 6 4 ...

FingerprintsVector;TopologicalPharmacophoreAtomTriplets:FixedSize:MinD
istance1:MaxDistance10;2692;OrderedNumericalValues;ValuesString;46 106
8 3 0 0 83 11 4 0 0 0 1 0 0 0 0 0 0 0 0 21 5 3 0 0 1 2 2 0 0 1 0 0 0
0 0 0 1 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 100 101 18 11 0 0 145 132 26
14 0 0 23 28 3 3 0 0 5 4 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 61 45 10 4 0
0 16 20 7 5 1 0 3 4 5 3 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 0 0 5 ...

FingerprintsVector;TopologicalPharmacophoreAtomTriplets:FixedSize:MinD
istance1:MaxDistance10;2692;OrderedNumericalValues;IDsAndValuesString;
Ar1-Ar1-Ar1 Ar1-Ar1-H1 Ar1-Ar1-HBA1 Ar1-Ar1-HBD1 Ar1-Ar1-NI1 Ar1-Ar1-P
I1 Ar1-H1-H1 Ar1-H1-HBA1 Ar1-H1-HBD1 Ar1-H1-NI1 Ar1-H1-PI1 Ar1-HBA1-HB
A1 Ar1-HBA1-HBD1 Ar1-HBA1-NI1 Ar1-HBA1-PI1 Ar1-HBD1-HBD1 Ar1-HBD1-...;
46 106 8 3 0 0 83 11 4 0 0 0 1 0 0 0 0 0 0 0 0 21 5 3 0 0 1 2 2 0 0 1
0 0 0 0 0 0 1 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 100 101 18 11 0 0 145
132 26 14 0 0 23 28 3 3 0 0 5 4 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 61 ...
```

METHODS

new

```
$TPATFP = new TopologicalPharmacophoreAtomTripletsFingerprints(
                                                    %NamesAndValues);
```

Using specified *TopologicalPharmacophoreAtomTripletsFingerprints* property names and values hash, new method creates a new object and returns a reference to newly created TopologicalPharmacophoreAtomTripletsFingerprints object. By default, the following properties are initialized:

```
Molecule = ''
Type = 'TopologicalPharmacophoreAtomTriplets'
MinDistance = 1
MaxDistance = 10
DistanceBinSize = 2
UseTriangleInequality = 1
AtomTypesToUse = ['HBD', 'HBA', 'PI', 'NI', 'H', 'Ar']
```

Examples:

```
$TPATFP = new TopologicalPharmacophoreAtomTripletsFingerprints(
                            'Molecule' => $Molecule);

$TPATFP = new TopologicalPharmacophoreAtomTripletsFingerprints(
                            'Molecule' => $Molecule,
                            'AtomTripletsSetSizeToUse' => 'ArbitrarySize';
                            'MinDistance' => 1,
                            'MaxDistance' => 10,
                            'DistanceBinSize' => 2,
                            'AtomTypesToUse' => ['HBD', 'HBA', 'PI', 'NI', 'H', 'Ar'],
```

```
                                'UseTriangleInequality' => 1);

    $TPATFP = new TopologicalPharmacophoreAtomTripletsFingerprints(
                                'Molecule' => $Molecule,
                                'AtomTripletsSetSizeToUse' => 'FixedSize';
                                'MinDistance' => 1,
                                'MaxDistance' => 10,
                                'DistanceBinSize' => 2,
                                'AtomTypesToUse' => ['HBD', 'HBA', 'PI', 'NI', 'H', 'Ar'],
                                'UseTriangleInequality' => 1);

    $TPATFP->GenerateFingerprints();
    print "$TPATFP\n";
```

GetDescription

```
    $Description = $TopologicalPharmacophoreAtomTripletsFP->GetDescription();
```

Returns a string containing description of topological pharmacophore atom triplets fingerprints.

GenerateFingerprints

```
    $TopologicalPharmacophoreAtomTripletsFP->GenerateFingerprints();
```

Generates topological pharmacophore atom triplets fingerprints and returns *TopologicalPharmacophoreAtomTripletsFP*.

GetAtomTripletIDs

```
    $AtomTripletsIDsRef = $TopologicalPharmacophoreATFP->GetAtomTripletIDs();
    @AtomTripletIDs = $TopologicalPharmacophoreATFP->GetAtomTripletIDs();
```

Returns atom triplet IDs corresponding to atom pairs count values in topological pharmacophore atom triplet fingerprints vector as an array or reference to an array.

AtomTripletsSetSizeToUse

```
    $TPAFP->AtomTripletsSetSizeToUse($Values);
```

Sets pharmacophore atom triplets set size to use for topological pharmacophore fingerprints generation and returns *TopologicalPharmacophoreAtomTripletsFingerprints*.

Possible values for pharmacophore atom triplets set size are: *ArbitrarySize, FizedSize*. Default value: *ArbitrarySize*.

For *ArbitrarySize* value of AtomTripletsSetSizeToUse, the fingerprint vector correspond to only those topological pharmacophore atom triplets which are present and have non-zero count. However, for *FixedSize* value of AtomTripletsSetSizeToUse, the fingerprint vector contains all possible valid topological pharmacophore atom triplets with both zero and non-zero count values.

SetAtomTypesToUse

```
    $TopologicalPharmacophoreAtomTripletsFP->SetAtomTypesToUse($ValuesRef);
    $TopologicalPharmacophoreAtomTripletsFP->SetAtomTypesToUse(@Values);
```

Sets pharmacophore atom types to use for topological pharmacophore fingerprints generation and returns *TopologicalPharmacophoreAtomTripletsFingerprints*.

Possible values for pharmacophore atom types are: *Ar, CA, H, HBA, HBD, Hal, NI, PI, RA*. Default value [Ref 71] : *HBD,HBA,PI,NI,H,Ar*.

The pharmacophore atom types abbreviations correspond to:

```
    HBD: HydrogenBondDonor
    HBA: HydrogenBondAcceptor
    PI :  PositivelyIonizable
    NI : NegativelyIonizable
    Ar : Aromatic
    Hal : Halogen
    H : Hydrophobic
    RA : RingAtom
    CA : ChainAtom
```

AtomTypes::FunctionalClassAtomTypes module is used to assign pharmacophore atom types. It uses following definitions [Ref 60-61, Ref 65-66]:

```
    HydrogenBondDonor: NH, NH2, OH
    HydrogenBondAcceptor: N[!H], O
    PositivelyIonizable: +, NH2
    NegativelyIonizable: -, C(=O)OH, S(=O)OH, P(=O)OH
```

SetDistanceBinSize

```
$TopologicalPharmacophoreAtomTripletsFP->SetDistanceBinSize($Value);
```

Sets distance bin size used to bin distances between atom pairs in atom triplets and returns *TopologicalPharmacophoreAtomTriplesFP*.

For default MinDistance and MaxDistance values of 1 and 10 with DistanceBinSize of 2 [Ref 70], the following 5 distance bins are generated:

```
[1, 2] [3, 4] [5, 6] [7, 8] [9 10]
```

The lower distance bound on the distance bin is uses to bin the distance between atom pairs in atom triplets. So in the previous example, atom pairs with distances 1 and 2 fall in first distance bin, atom pairs with distances 3 and 4 fall in second distance bin and so on.

In order to distribute distance bins of equal size, the last bin is allowed to go past MaxDistance by up to distance bin size. For example, MinDistance and MaxDistance values of 2 and 10 with DistanceBinSize of 2 generates the following 6 distance bins:

```
[2, 3] [4, 5] [6, 7] [8, 9] [10 11]
```

SetMaxDistance

```
$TopologicalPharmacophoreAtomTriplesFP->SetMaxDistance($Value);
```

Sets maximum bond distance between atom pairs corresponding to atom triplets for generating topological pharmacophore atom triplets fingerprints and returns *TopologicalPharmacophoreAtomTriplesFP*.

SetMinDistance

```
$TopologicalPharmacophoreAtomTriplesFP->SetMinDistance($Value);
```

Sets minimum bond distance between atom pairs corresponding to atom triplets for generating topological pharmacophore atom triplets fingerprints and returns *TopologicalPharmacophoreAtomTriplesFP*.

StringifyTopologicalPharmacophoreAtomTripletsFingerprints

```
$String = $TopologicalPharmacophoreAtomTripletsFingerprints->
                StringifyTopologicalPharmacophoreAtomTripletsFingerprints();
```

Returns a string containing information about *TopologicalPharmacophoreAtomTripletsFingerprints* object.

AUTHOR

Manish Sud <msud@san.rr.com>

SEE ALSO

Fingerprints.pm, FingerprintsStringUtil.pm, AtomNeighborhoodsFingerprints.pm, AtomTypesFingerprints.pm, EStateIndiciesFingerprints.pm, ExtendedConnectivityFingerprints.pm, MACCSKeys.pm, PathLengthFingerprints.pm, TopologicalAtomPairsFingerprints.pm, TopologicalAtomTripletsFingerprints.pm, TopologicalAtomTorsionsFingerprints.pm, TopologicalPharmacophoreAtomPairsFingerprints.pm,

COPYRIGHT