

NAME

TopologicalAtomTripletsFingerprints

SYNOPSIS

```
use Fingerprints::TopologicalAtomTripletsFingerprints;

use Fingerprints::TopologicalAtomTripletsFingerprints qw(:all);
```

DESCRIPTION

TopologicalAtomTripletsFingerprints [Ref 57, Ref 59, Ref 72] class provides the following methods:

new, GenerateFingerprints, GetAtomTripletIDs, GetDescription, SetAtomIdentifierType, SetAtomicInvariantsToUse, SetFunctionalClassesToUse, SetMaxDistance, SetMinDistance, StringifyTopologicalAtomTripletsFingerprints

TopologicalAtomTripletsFingerprints is derived from Fingerprints class which in turn is derived from ObjectProperty base class that provides methods not explicitly defined in TopologicalAtomTripletsFingerprints, 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>();
```

The current release of MayaChemTools supports generation of TopologicalAtomTripletsFingerprints corresponding to following AtomIdentifierTypes:

```
AtomicInvariantsAtomTypes, DREIDINGAtomTypes, EStateAtomTypes,
FunctionalClassAtomTypes, MMFF94AtomTypes, SLogPAtomTypes,
SYBYLAtomTypes, TPSAAtomTypes, UFFAtomTypes
```

Based on the values specified for AtomIdentifierType along with other specified parameters such as AtomicInvariantsToUse and FunctionalClassesToUse, initial atom types are assigned to all non-hydrogen atoms in a molecule. Using the distance matrix for the molecule and initial atom types assigned to non-hydrogen atoms, all unique atom triplets within MinDistance and MaxDistance are identified and counted. An atom triplet identifier is generated for each unique atom triplet; the format of atom triplet identifier is:

```
<ATx>-Dyz-<ATy>-Dxz-<ATz>-Dxy
```

```
ATx, ATy, ATz: Atom types assigned to atom x, atom y, and atom z
Dxy: Distance between atom x and atom y
Dxz: Distance between atom x and atom z
Dyz: Distance between atom y and atom z
```

```
where <AT1>-D23 <= <AT2>-D13 <= <AT3>-D12
```

The atom triplet identifiers for all unique atom triplets corresponding to non-hydrogen atoms constitute topological atom triplets fingerprints of the molecule.

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

```
FingerprintsVector;TopologicalAtomTriplets:AtomicInvariantsAtomTypes:M
inDistance1:MaxDistance10;3096;NumericalValues;IDsAndValuesString;C.X1
.B01.H3-D1-C.X1.B01.H3-D1-C.X3.B03.H1-D2 C.X1.B01.H3-D1-C.X2.B02.H2-D1
0-C.X3.B04-D9 C.X1.B01.H3-D1-C.X2.B02.H2-D3-N.X3.B03-D4 C.X1.B01.H3-D1
-C.X2.B02.H2-D4-C.X2.B02.H2-D5 C.X1.B01.H3-D1-C.X2.B02.H2-D6-C.X3...;
1 2 2 2 2 2 2 8 8 4 8 4 4 2 2 2 2 4 2 2 2 2 1 2 2 4 4 4 2 2
2 4 4 4 8 4 4 2 4 4 4 2 4 4 2 2 2 2 2 2 1 2 2 2 2 2 2 2 2 8...
```

```
FingerprintsVector;TopologicalAtomTriplets:AtomicInvariantsAtomTypes:M
inDistance1:MaxDistance10;3096;NumericalValues;IDsAndValuesPairsString
;C.X1.B01.H3-D1-C.X1.B01.H3-D1-C.X3.B03.H1-D2 1 C.X1.B01.H3-D1-C.X2.B0
2.H2-D10-C.X3.B04-D9 2 C.X1.B01.H3-D1-C.X2.B02.H2-D3-N.X3.B03-D4 2 C.X
1.B01.H3-D1-C.X2.B02.H2-D4-C.X2.B02.H2-D5 2 C.X1.B01.H3-D1-C.X2.B02.H2
-D6-C.X3.B03.H1-D5 2 C.X1.B01.H3-D1-C.X2.B02.H2-D6-C.X3.B03.H1-D7 2...
```

```
FingerprintsVector;TopologicalAtomTriplets:DREIDINGAtomTypes:MinDistance1:MaxDistance10;2377;NumericalValues;IDsAndValuesString;C_2-D1-C_2-D9-C_3-D10 C_2-D1-C_2-D9-C_R-D10 C_2-D1-C_3-D1-C_3-D2 C_2-D1-C_3-D10-C_3-D9 C_2-D1-C_3-D2-C_3-D3 C_2-D1-C_3-D2-C_R-D3 C_2-D1-C_3-D3-C_3-D4 C_2-D1-C_3-D3-N_R-D4 C_2-D1-C_3-D3-O_3-D2 C_2-D1-C_3-D4-C_3-D5 C_2-D...;1 1 1 2 1 1 3 1 1 2 2 1 1 1 1 1 1 1 2 1 3 4 5 1 1 6 4 2 2 3 1 1 1 2 2 1 2 1 1 2 2 2 1 2 1 2 1 1 3 3 2 6 4 2 1 1 1 2 2 1 1 1 1 1 1 1 1 1 1 1...
```

```
FingerprintsVector;TopologicalAtomTriplets:EStateAtomTypes:MinDistance1:MaxDistance10;3298;NumericalValues;IDsAndValuesString;aaCH-D1-aaCH-D1-aaCH-D2 aaCH-D1-aaCH-D1-aasC-D2 aaCH-D1-aaCH-D10-aaCH-D9 aaCH-D1-aaCH-D10-aasC-D9 aaCH-D1-aaCH-D2-aaCH-D3 aaCH-D1-aaCH-D2-aasC-D1 aaCH-D1-aaCH-D2-aasC-D3 aaCH-D1-aaCH-D3-aasC-D2 aaCH-D1-aaCH-D4-aasC-D5 aa...;6 4 24 4 16 8 8 4 8 8 8 12 10 14 4 16 24 4 12 2 2 4 1 10 2 2 15 2 2 2 2 2 14 4 2 2 2 1 2 10 2 2 4 1 2 4 8 3 3 3 4 6 4 2 2 3 3 1 1 1 2 1 2 2 4 2 3 2 1 2 4 5 3 2 2 1 2 4 3 2 8 12 6 2 2 4 4 7 1 4 2 4 2 2 2 ...
```

```
FingerprintsVector;TopologicalAtomTriplets:FunctionalClassAtomTypes:MinDistance1:MaxDistance10;2182;NumericalValues;IDsAndValuesString;Ar-D1-Ar-D1-Ar-D2 Ar-D1-Ar-D1-Ar.HBA-D2 Ar-D1-Ar-D10-Ar-D9 Ar-D1-Ar-D10-Hal-D9 Ar-D1-Ar-D2-Ar-D2 Ar-D1-Ar-D2-Ar-D3 Ar-D1-Ar-D2-Ar.HBA-D1 Ar-D1-Ar-D2-Ar.HBA-D2 Ar-D1-Ar-D2-Ar.HBA-D3 Ar-D1-Ar-D2-HBD-D1 Ar-D1-Ar-D2...;27 1 32 2 2 63 3 2 1 2 1 2 3 1 1 40 3 1 2 2 2 2 4 2 2 47 4 2 2 1 2 1 5 2 2 51 4 3 1 3 1 9 1 1 50 3 3 4 1 9 50 2 2 3 3 5 45 1 1 1 2 1 2 2 3 3 4 4 3 2 1 1 3 4 5 5 3 1 2 3 2 3 5 7 2 7 3 7 1 1 2 2 2 2 3 1 4 3 1 2...
```

```
FingerprintsVector;TopologicalAtomTriplets:MMFF94AtomTypes:MinDistance1:MaxDistance10;2966;NumericalValues;IDsAndValuesString;C5A-D1-C5A-D1-N5-D2 C5A-D1-C5A-D2-C5B-D2 C5A-D1-C5A-D3-CB-D2 C5A-D1-C5A-D3-CR-D2 C5A-D1-C5B-D1-C5B-D2 C5A-D1-C5B-D2-C=ON-D1 C5A-D1-C5B-D2-CB-D1 C5A-D1-C5B-D3-C=ON-D2 C5A-D1-C5B-D3-CB-D2 C5A-D1-C=ON-D3-NC=O-D2 C5A-D1-C=ON-D3-O=CN-D2 C5A-D1-C=ON-D4-NC=O-D3 C5A-D1-C=ON-D4-O=CN-D3 C5A-D1-CB-D1...
```

```
FingerprintsVector;TopologicalAtomTriplets:SLogPAtomTypes:MinDistance1:MaxDistance10;3710;NumericalValues;IDsAndValuesString;C1-D1-C1-D1-C11-D2 C1-D1-C1-D1-CS-D2 C1-D1-C1-D10-C5-D9 C1-D1-C1-D3-C10-D2 C1-D1-C1-D3-C5-D2 C1-D1-C1-D3-CS-D2 C1-D1-C1-D3-CS-D4 C1-D1-C1-D4-C10-D5 C1-D1-C1-D4-C11-D5 C1-D1-C1-D5-C10-D4 C1-D1-C1-D5-C5-D4 C1-D1-C1-D6-C11-D7 C1-D1-C1-D6-CS-D5 C1-D1-C1-D6-CS-D7 C1-D1-C1-D8-C11-D9 C1-D1-C1-D8-CS...
```

```
FingerprintsVector;TopologicalAtomTriplets:SYBYLAtomTypes:MinDistance1:MaxDistance10;2332;NumericalValues;IDsAndValuesString;C.2-D1-C.2-D9-C.3-D10 C.2-D1-C.2-D9-C.ar-D10 C.2-D1-C.3-D1-C.3-D2 C.2-D1-C.3-D10-C.3-D9 C.2-D1-C.3-D2-C.3-D3 C.2-D1-C.3-D2-C.ar-D3 C.2-D1-C.3-D3-C.3-D4 C.2-D1-C.3-D3-N.ar-D4 C.2-D1-C.3-D3-O.3-D2 C.2-D1-C.3-D4-C.3-D5 C.2-D1-C.3-D5-C.3-D6 C.2-D1-C.3-D5-O.3-D4 C.2-D1-C.3-D6-C.3-D7 C.2-D1-C.3-D7...
```

```
FingerprintsVector;TopologicalAtomTriplets:TPSAAAtomTypes:MinDistance1:MaxDistance10;1007;NumericalValues;IDsAndValuesString;N21-D1-N7-D3-None-D4 N21-D1-N7-D5-None-D4 N21-D1-None-D1-None-D2 N21-D1-None-D2-None-D2 N21-D1-None-D2-None-D3 N21-D1-None-D3-None-D4 N21-D1-None-D4-None-D5 N21-D1-None-D4-O3-D3 N21-D1-None-D4-O4-D3 N21-D1-None-D5-None-D6 N21-D1-None-D6-None-D7 N21-D1-None-D6-O4-D5 N21-D1-None-D7-None-D8 N21...
```

```
FingerprintsVector;TopologicalAtomTriplets:UFFAtomTypes:MinDistance1:MaxDistance10;2377;NumericalValues;IDsAndValuesString;C_2-D1-C_2-D9-C_3-D10 C_2-D1-C_2-D9-C_R-D10 C_2-D1-C_3-D1-C_3-D2 C_2-D1-C_3-D10-C_3-D9 C_2-D1-C_3-D2-C_3-D3 C_2-D1-C_3-D2-C_R-D3 C_2-D1-C_3-D3-C_3-D4 C_2-D1-C_3-D3-N_R-D4 C_2-D1-C_3-D3-O_3-D2 C_2-D1-C_3-D4-C_3-D5 C_2-D1-C_3-D5-C_3-D6 C_2-D1-C_3-D5-O_3-D4 C_2-D1-C_3-D6-C_3-D7 C_2-D1-C_3-D7-C_3...
```

METHODS

new

```
$NewTopologicalAtomTripletsFingerprints = new TopologicalAtomTripletsFingerprints(%NamesAndValues);
```

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

```
Molecule = ''
Type = 'TopologicalAtomTriplets'
MinDistance = 1
MaxDistance = 10
UseTriangleInequality = 1
AtomIdentifierType = ''
AtomicInvariantsToUse = ['AS', 'X', 'BO', 'H', 'FC']
FunctionalClassesToUse = ['HBD', 'HBA', 'PI', 'NI', 'Ar', 'Hal']
```

Examples:

```
$TopologicalAtomTripletsFingerprints = new TopologicalAtomTripletsFingerprints(
    'Molecule' => $Molecule,
    'AtomIdentifierType' =>
        'AtomicInvariantsAtomTypes');

$TopologicalAtomTripletsFingerprints = new TopologicalAtomTripletsFingerprints(
    'Molecule' => $Molecule,
    'MinDistance' => 1,
    'MaxDistance' => 10,
    'AtomIdentifierType' =>
        'AtomicInvariantsAtomTypes',
    'AtomicInvariantsToUse' =>
        ['AS', 'X', 'BO', 'H', 'FC']);

$TopologicalAtomTripletsFingerprints = new TopologicalAtomTripletsFingerprints(
    'Molecule' => $Molecule,
    'AtomIdentifierType' =>
        'DREIDINGAtomTypes');

$TopologicalAtomTripletsFingerprints = new TopologicalAtomTripletsFingerprints(
    'Molecule' => $Molecule,
    'AtomIdentifierType' =>
        'MMFF94AtomTypes');

$TopologicalAtomTripletsFingerprints = new TopologicalAtomTripletsFingerprints(
    'Molecule' => $Molecule,
    'AtomIdentifierType' =>
        'TPSAAAtomTypes');

$TopologicalAtomTripletsFingerprints = new TopologicalAtomTripletsFingerprints(
    'Molecule' => $Molecule,
    'MinDistance' => 1,
    'MaxDistance' => 10,
    'AtomIdentifierType' =>
        'FunctionalClassAtomTypes',
    'FunctionalClassesToUse' =>
        ['HBD', 'HBA', 'PI', 'NI', 'Ar', 'Hal']);

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

GetDescription

```
$Return = $TopologicalAtomTripletsFingerprints->GetDescription();
```

Returns a string containing description of topological atom triplets fingerprints.

GenerateFingerprints

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

Generates topological atom triplets fingerprints and returns *TopologicalAtomTripletsFingerprints*.

GetAtomTripletIDs

```
$AtomTripletIDsRef = $TopologicalAtomTripletsFingerprints->GetAtomTripletIDs();
@AtomTripletIDs = $TopologicalAtomTripletsFingerprints->GetAtomTripletIDs();
```

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

SetAtomIdentifierType

```
$TopologicalAtomTripletsFingerprints->SetAtomIdentifierType($IdentifierType);
```

Sets atom *IdentifierType* to use during atom triplets fingerprints generation and returns *TopologicalAtomTripletsFingerprints*.

Possible values: *AtomicInvariantsAtomTypes*, *DREIDINGAtomTypes*, *EStateAtomTypes*, *FunctionalClassAtomTypes*, *MMFF94AtomTypes*, *SLogPAtomTypes*, *SYBYLAtomTypes*, *TPSAAAtomTypes*, *UFFAtomTypes*.

SetAtomicInvariantsToUse

```
$TopologicalAtomTripletsFingerprints->SetAtomicInvariantsToUse($ValuesRef);
$TopologicalAtomTripletsFingerprints->SetAtomicInvariantsToUse(@Values);
```

Sets atomic invariants to use during *AtomicInvariantsAtomTypes* value of *AtomIdentifierType* for topological atom triplets fingerprints generation and returns *TopologicalAtomTripletsFingerprints*.

Possible values for atomic invariants are: *AS*, *X*, *BO*, *LBO*, *SB*, *DB*, *TB*, *H*, *Ar*, *RA*, *FC*, *MN*, *SM*. Default value: *AS,X,BO,H,FC*.

The atomic invariants abbreviations correspond to:

AS = Atom symbol corresponding to element symbol

X<n> = Number of non-hydrogen atom neighbors or heavy atoms
BO<n> = Sum of bond orders to non-hydrogen atom neighbors or heavy atoms
LBO<n> = Largest bond order of non-hydrogen atom neighbors or heavy atoms
SB<n> = Number of single bonds to non-hydrogen atom neighbors or heavy atoms
DB<n> = Number of double bonds to non-hydrogen atom neighbors or heavy atoms
TB<n> = Number of triple bonds to non-hydrogen atom neighbors or heavy atoms
H<n> = Number of implicit and explicit hydrogens for atom
Ar = Aromatic annotation indicating whether atom is aromatic
RA = Ring atom annotation indicating whether atom is a ring
FC<+n/-n> = Formal charge assigned to atom
MN<n> = Mass number indicating isotope other than most abundant isotope
SM<n> = Spin multiplicity of atom. Possible values: 1 (singlet), 2 (doublet) or 3 (triplet)

Atom type generated by *AtomTypes::AtomicInvariantsAtomTypes* class corresponds to:

```
AS.X<n>.BO<n>.LBO<n>.<SB><n>.<DB><n>.<TB><n>.H<n>.Ar.RA.FC<+n/-n>.MN<n>.SM<n>
```

Except for *AS* which is a required atomic invariant in atom types, all other atomic invariants are optional. Atom type specification doesn't include atomic invariants with zero or undefined values.

In addition to usage of abbreviations for specifying atomic invariants, the following descriptive words are also allowed:

X : NumOfNonHydrogenAtomNeighbors or NumOfHeavyAtomNeighbors
BO : SumOfBondOrdersToNonHydrogenAtoms or SumOfBondOrdersToHeavyAtoms
LBO : LargestBondOrderToNonHydrogenAtoms or LargestBondOrderToHeavyAtoms
SB : NumOfSingleBondsToNonHydrogenAtoms or NumOfSingleBondsToHeavyAtoms
DB : NumOfDoubleBondsToNonHydrogenAtoms or NumOfDoubleBondsToHeavyAtoms
TB : NumOfTripleBondsToNonHydrogenAtoms or NumOfTripleBondsToHeavyAtoms
H : NumOfImplicitAndExplicitHydrogens
Ar : Aromatic
RA : RingAtom
FC : FormalCharge
MN : MassNumber
SM : SpinMultiplicity

AtomTypes::AtomicInvariantsAtomTypes module is used to assign atomic invariant atom types.

SetFunctionalClassesToUse

```
$TopologicalTripletsFingerprints->SetFunctionalClassesToUse($ValuesRef);
$TopologicalTripletsFingerprints->SetFunctionalClassesToUse(@Values);
```

Sets functional classes invariants to use during *FunctionalClassAtomTypes* value of *AtomIdentifierType* for topological atom triplets fingerprints generation and returns *TopologicalAtomTripletsFingerprints*.

Possible values for atom functional classes are: *Ar, CA, H, HBA, HBD, Hal, NI, PI, RA*. Default value [Ref 24]: *HBD,HBA,PI,NI,Ar,Hal*.

The functional class abbreviations correspond to:

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

Functional class atom type specification for an atom corresponds to:

```
Ar.CA.H.HBA.HBD.Hal.NI.PI.RA or None
```

AtomTypes::FunctionalClassAtomTypes module is used to assign functional class 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
```

SetMaxDistance

```
$TopologicalAtomTripletsFingerprints->SetMaxDistance($Distance);
```

Sets maximum distance to use during topological atom triplets fingerprints generation and returns *TopologicalAtomTripletsFingerprints*.

SetMinDistance

```
$TopologicalAtomTripletsFingerprints->SetMinDistance($Distance);
```

Sets minimum distance to use during topological atom triplets fingerprints generation and returns *TopologicalAtomTripletsFingerprints*.

StringifyTopologicalAtomTripletsFingerprints

```
$String = $TopologicalAtomTripletsFingerprints->
    StringifyTopologicalAtomTripletsFingerprints();
```

Returns a string containing information about *TopologicalAtomTripletsFingerprints* object.

AUTHOR

Manish Sud <msud@san.rr.com>

SEE ALSO

Fingerprints.pm, FingerprintsStringUtil.pm, AtomNeighborhoodsFingerprints.pm, AtomTypesFingerprints.pm, EStateIndicesFingerprints.pm, ExtendedConnectivityFingerprints.pm, MACCSKeys.pm, PathLengthFingerprints.pm, TopologicalAtomPairsFingerprints.pm, TopologicalAtomTorsionsFingerprints.pm, TopologicalPharmacophoreAtomPairsFingerprints.pm, TopologicalPharmacophoreAtomTripletsFingerprints.pm

COPYRIGHT

Copyright (C) 2020 Manish Sud. All rights reserved.

This file is part of MayaChemTools.

MayaChemTools is free software; you can redistribute it and/or modify it under the terms of the GNU Lesser General Public License as published by the Free Software Foundation; either version 3 of the License, or (at your option) any later version.