RESEARCH

Substructure-based Neural Machine Translation for Retrosynthetic Prediction

Umit V. Ucak[1], Taek Kang[2], Junsu Ko[3*] and Juyong Lee[1*]

[*]Correspondence:
junsuko@arontier.co;
juyong.lee@kangwon.ac.kr
[3]Arontier co., Seoul, South Korea
[1]Division of Chemistry and Biochemistry, Department of Chemistry, Kangwon National University, Chuncheon, South Korea
Full list of author information is available at the end of the article

Abstract

Keywords: retrosynthesis planning; machine neural translation; seq-to-seq; attention

Additional Files as Figures.

Please find the supporting materials as **figures** within the "Additional Files" section of the BMC article.

Author details
[1]Division of Chemistry and Biochemistry, Department of Chemistry, Kangwon National University, Chuncheon, South Korea. [2]Center for Neuro-Medicine, Korea Institute of Science and Technology, Seoul, South Korea. [3]Arontier co., Seoul, South Korea.

References

Additional Files

Additional File 3 : Figure S2

File name : Supplementary Figure S2
Title of data : 1-bit keys
File format : Standard Latex figure, formatted as PNG.
Description of data : Examples to molecules that are represented with only one bit.

Lettered : MACCS : SMARTS Key	Key	Pattern	SMARTS Visual	Corresponding Molecules (SMILES)
'Wz' : '17' : '[#6]#[#6]'				'C#C', '[C-]#[C-]', 'C#CBr'
'Mx' : '88' : '[#16]'				'S', '[S-2]', '[SH-]'
'dz' : '99' : '[#6] = [#6]'				'FC=C[SnH3]', 'BrC=CBr', 'Cl/C=C\\Cl'
'ox' : '106' : '[!#6;!#1] ~ *(~ [!#6;!#1]) ~ [!#6;!#1]'				'ClC(Cl)Cl', 'BrC(Br)Br', 'FC(F)Cl', 'FC(Cl)Cl', '[SiH3]C(Cl)Cl', 'IC(I)I'
'ix' : '124' : '[!#6;!#1] ~ [!#6;!#1]'				'ClCl', 'BrBr', 'II', 'FF', '[Se-][Se-]', 'ClI', 'BrI', 'ClBr'
's' : '160' : '[C;H3,H4]'				'C', 'CC(Cl)Cl', 'CC(Br)Br', 'CC(F)Br', 'CC(Cl)Br'
't' : '164' : '[#8]'				'[O-2]'

Figure S2: Molecules that are represented with only one bit in the double reactant dataset are given as SMILES strings.