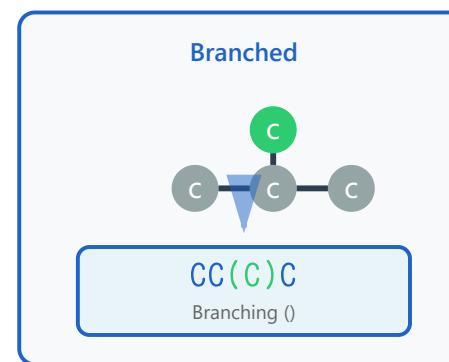
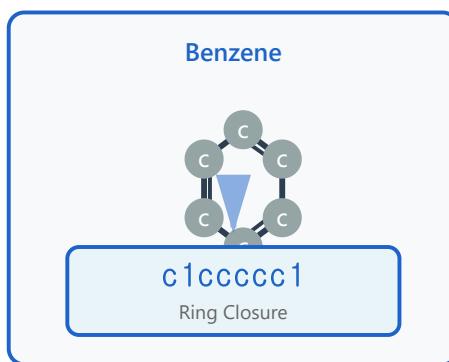
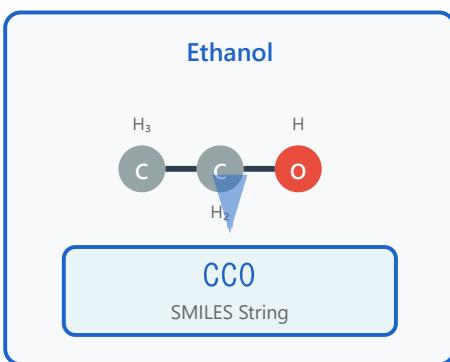


SMILES Notation



Syntax rules

String-based molecular encoding

Canonical SMILES

Unique molecular representation

SMARTS patterns

Substructure search patterns

Tokenization

Breaking into meaningful units

Augmentation strategies

Data augmentation techniques