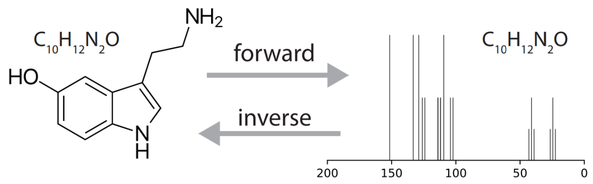
INTERNATIONAL INSTITUTE OF INFORMATION TECHNOLOGY
HYDERABAD
R&D SHOWCASE 2021
Technology, Social Impact

Inverse Problem of determining the Molecular Structure from NMR Spectra

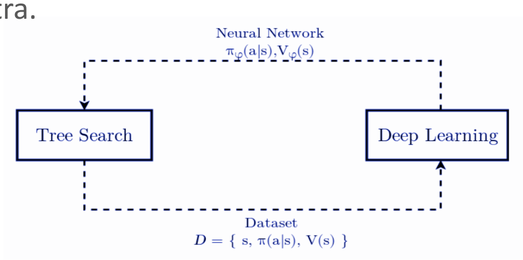
ABSTRACTS

Nuclear Magnetic Resonance is an important tool in identifying molecules in a sample. Despite having a large database of previously observed NMR Spectra, we have only discovered a tiny fraction of the actual chemical space. With advent of machine learning and deep learning methods, this calls for a way to be capable of finding the nature and structure of sample from its NMR Spectra.



OBJECTIVE

The aim of this project is to device an efficient and dependable framework which can understand the mapping between chemical space and NMR Spectroscopy. And hence can navigate the chemical space to correctly determine the molecular structure of a substance given its NMR Spectra.

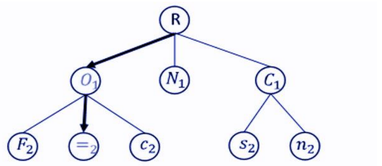


METHOD

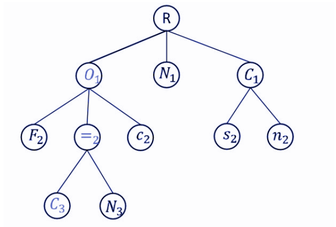
This project builds upon a molecule from scratch by making actions (adding a bond) at each stage(Current Molecule). Action at each stage is determined by a run of Monte Carlo Tree Search which is guided by two neural networks – Prior Network and Value Network. The motive of the tree is to try to balance exploitation vs exploration of prior knowledge by updating the UCT Values after each selection.

$$UCT = Q_{s,a} + c \cdot p_s \cdot \sqrt{\frac{N_s}{n_{s,a}}}$$

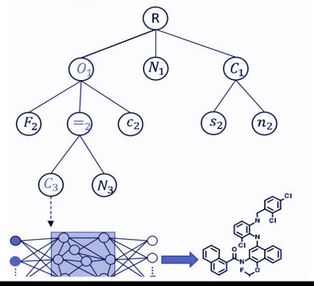
Selection



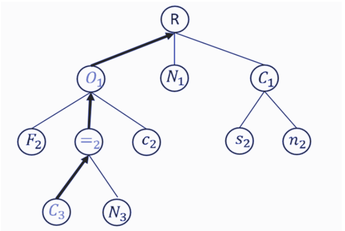
Expansion



Roll-out



Backpropagate



References:

1. Deep imitation learning for molecular inverse problems E Jonas - Advances in Neural Information Processing Systems, 2019
2. ChemTS: An Efficient Python Library for de novo Molecular Generation, https://arxiv.org/abs/1710.00616

Authors: Bhuvanesh Sridharan, Sarvesh Mehta, Yashaswi Pathak and Prof. Deva Priyakumar

Research Center Name: Centre for Computational Natural Science and Bioinformatics