Task 2: Making the dictionary

How to make a dictionary that is versatile, easy to access, and is analyzable:

Idea: Save each key into a value that is a list.

Remember that from task 1, the list is saved as: [index, string element, int ring_status, array of outer connection, array of inner connection, bool isedge]

isedge is not very important, so we will ignore that for now

ring status should be the first factor: there will be 7 cases: benzene, benzene border, 6-ring, 6-ring border, 5-ring, 5-ring border, no ring, denoted from 0 to 6

elements and connection should be considered together. Note that for most cases we do not need to worry about sidechains, we will only need another bollean(int 0-1) that basically ask is there a sidechain: 0: no, 1: yes

How to best express this: for the fully ring case, we can just express them as a string combining the atom and the connection just like smiles

For the border case: we will have a string represent the inner and the outer will be put in ()

For the non-ring case: we will have the longest path outside and the sidechain will be put in () (for now) A better solution: looking at the martini bead, the bead can have sidechains but it seems like the bead can only have sidechains from 1 node, we can use this property and just take the center node and put every other section in ()

The next section is the amount of extra connection, this can be done by tracing the connectivity matrix and find out how many unaccounted connection does it have (this may not be helpful but it is good to add as an argument)

we can add another section that specify what bead this bead (usually) connects to for even further breakdown if there are still confusion

Current table:

def get_m3_dict(): martini_dict = {}

```
#Break the bead types into section
#Form: [section, sidechains?, representation(if sidechain only the
#center node is outside and each direction is another ()), ]
#Section 1: Benzene nodes
martini_dict["TN6a"] = [1, 0, "CN", 2, ""]
martini_dict["TC5"] = [1, 0, "CC", 2, ""]
#Section 2: Benzene border nodes
martini_dict["SN6"] = [2, 0, "CC(O)", 2, ""]
martini_dict["SN6d"] = [2, 0, "CC(N)", 2, ""]
```

```
martini_dict["SC6"] = [2, 0, "CC(S)", 2, ""]
martini_dict["SX3"] = [2, 0, "CC(Cl)", 2, ""]
martini_dict["X1"] = [2, 0, "CC(I)", 2, ""]
martini_dict["SC4"] = [2, 0, "CC(C)", 2, ""]
martini_dict["TN6"] = [2, 0, "C(O)", 2, ""]
martini_dict["TC4"] = [2, 0, "C(C)", 2, ""]
martini_dict["TN6a+"] = [2, 0, "C(=O)", 2, ""]
martini_dict["SN2a"] = [2, 0, "C(OC)", 2, ""]
#Section 3: Non-benzene 6-ring nodes
martini_dict["SN4a"] = [3, 0, "COC", 2, "SC3"]
martini_dict["SN3a"] = [3, 0, "COC", 2, "SN3a"]
martini_dict["SC3"] = [3, 0, "CCC", 2, ""]
#Section 4: Non-benzene 6-ring border nodes
#Section 5: 5-ring nodes
martini_dict["TN6d"] = [5, 0, "NH", 2, ""]
martini_dict["SN5a"] = [5, 0, "OCO", 2, ""]
martini_dict["SC6+"] = [5, 0, "CSC", 2, ""]
martini_dict["TN4a"] = [5, 0, "OC", 2, ""]
martini_dict["TC3"] = [5, 0, "CC", 2, ""]
martini_dict["TN2a"] = [5, 0, "O", 2, ""]
martini_dict["TC6"] = [5, 0, "S", 2, ""]
#Section 6: 5-ring border nodes
martini_dict["SN6+"] = [6, 0, "CC(O)", 2, ""]
martini_dict["SC3+"] = [6, 0, "CC(C)", 2, ""]
martini_dict["TN1"] = [6, 0, "N(C)", 2, ""]
#Section 7: Non-ring nodes
martini_dict["P2"] = [7, 1, "C(O)(=O)(CC)", 0, ""]
martini_dict["SP2"] = [7, 1, "C(O)(=O)(C)", 0, ""]
martini_dict["P1"] = [7, 1, "C(O)(C)(C)", 0, ""]
martini_dict["N4a+"] = [7, 1, "C(C)(=O)(OC)", 0, ""]
martini_dict["SX4e"] = [7, 1, "C(F)(F)(F)", 1, ""]
martini_dict["SX2"] = [7, 1, "C(Cl)(Cl)(Cl)", 0, ""]
martini_dict["SN3r"] = [7, 0, "COC", 0, ""]
martini_dict["TP1d"] = [7, 0, "CO", 1, "SX4e"]
martini_dict["TN2a"] = [7, 0, "CO", 1, "TC5"]
martini_dict["N6a"] = [7, 0, "CCC=O", 1, "C1"]
martini_dict["N4a"] = [7, 0, "COC=O", 1, ""]
martini_dict["SN4a"] = [7, 0, "CC=O", 1, ""]
martini_dict["TN4a+"] = [7, 0, "N#C", 1, ""]
martini_dict["TN4a++"] = [7, 0, "C=O", 1, ""]
martini_dict["SN3a"] = [7, 0, "ON=O", 1, ""]
martini_dict["SN1"] = [7, 0, "CNC", 1, ""]
martini_dict["TC5+"] = [7, 0, "C=C", 1, ""]
```

```
martini_dict["TC3+"] = [7, 0, "CC", 1, ""]
martini_dict["C4"] = [7, 0, "CC=CC", 0, ""]
martini_dict["C1"] = [7, 0, "CCCC", 0, ""]
martini_dict["SC2"] = [7, 0, "CCC", 1, ""]
#Section 8: node that is in 2 benzene rings:
martini_dict["TC5e"] = [8, 0, "CC", 4, ""]

return martini_dict
```