execfile(filename, namespace)

File "C:\CS\Anaconda\lib\site-packages\spyder\utils\site\sitecustomize.py", line 101, in execfile

exec(compile(f.read(), filename, 'exec'), namespace)

File "C:/Users/njayj/OneDrive/Documents/CS/Bio/debruijn.py", line 17, in <module>

nodes, edges = de\_bruijn("ACGCGTCG", 3)

NameError: name 'de\_bruijn' is not defined

runfile('C:/Users/njayj/OneDrive/Documents/CS/Bio/debruijn.py', wdir='C:/Users/njayj/OneDrive/Documents/CS/Bio')

nodes

Out[9]: {'AC', 'CG', 'GC', 'GT', 'TC'}

edges

Out[10]:

[('AC', 'CG'),

('CG', 'GC'),

('GC', 'CG'),

('CG', 'GT'),

('GT', 'TC'),

('TC', 'CG')]

runfile('C:/Users/njayj/OneDrive/Documents/CS/Bio/debruijn.py', wdir='C:/Users/njayj/OneDrive/Documents/CS/Bio')

load\_ext gvmagic

Traceback (most recent call last):

File "<ipython-input-12-f960a0c125dc>", line 1, in <module>

get\_ipython().magic('load\_ext gvmagic')

File "C:\CS\Anaconda\lib\site-packages\IPython\core\interactiveshell.py", line 2146, in magic

return self.run\_line\_magic(magic\_name, magic\_arg\_s)

File "C:\CS\Anaconda\lib\site-packages\IPython\core\interactiveshell.py", line 2067, in run\_line\_magic

result = fn(\*args,\*\*kwargs)

File "<decorator-gen-66>", line 2, in load\_ext

File "C:\CS\Anaconda\lib\site-packages\IPython\core\magic.py", line 187, in <lambda>

call = lambda f, \*a, \*\*k: f(\*a, \*\*k)

File "C:\CS\Anaconda\lib\site-packages\IPython\core\magics\extension.py", line 33, in load\_ext

res = self.shell.extension\_manager.load\_extension(module\_str)

File "C:\CS\Anaconda\lib\site-packages\IPython\core\extensions.py", line 85, in load\_extension

mod = import\_module(module\_str)

File "C:\CS\Anaconda\lib\importlib\\_\_init\_\_.py", line 126, in import\_module

return \_bootstrap.\_gcd\_import(name[level:], package, level)

File "<frozen importlib.\_bootstrap>", line 994, in \_gcd\_import

File "<frozen importlib.\_bootstrap>", line 971, in \_find\_and\_load

File "<frozen importlib.\_bootstrap>", line 953, in \_find\_and\_load\_unlocked

ModuleNotFoundError: No module named 'gvmagic'

dotstr visualize\_de\_bruijn("ACGCGTCG", 3)

File "<ipython-input-13-3433cc4cf091>", line 1

dotstr visualize\_de\_bruijn("ACGCGTCG", 3)

^

SyntaxError: invalid syntax

a = dotstr visualize\_de\_bruijn("ACGCGTCG", 3)

File "<ipython-input-14-b35b297f7722>", line 1

a = dotstr visualize\_de\_bruijn("ACGCGTCG", 3)

^

SyntaxError: invalid syntax

a = dotstr visualize\_de\_bruijn("ACGCGTCG", 3)

File "<ipython-input-15-b35b297f7722>", line 1

a = dotstr visualize\_de\_bruijn("ACGCGTCG", 3)

^

SyntaxError: invalid syntax

dotstr = visualize\_de\_bruijn("ACGCGTCG", 3)

Traceback (most recent call last):

File "<ipython-input-16-ef9cf1aad697>", line 1, in <module>

dotstr = visualize\_de\_bruijn("ACGCGTCG", 3)

File "C:/Users/njayj/OneDrive/Documents/CS/Bio/debruijn.py", line 21, in visualize\_de\_bruijn

nodes, edges = de\_bruijn\_ize(st, k)

NameError: name 'de\_bruijn\_ize' is not defined

runfile('C:/Users/njayj/OneDrive/Documents/CS/Bio/debruijn.py', wdir='C:/Users/njayj/OneDrive/Documents/CS/Bio')

dotstr = visualize\_de\_bruijn("ACGCGTCG", 3)

dotstr

Out[19]: 'digraph "DeBruijn graph" {\n AC [label="AC"] ;\n GC [label="GC"] ;\n TC [label="TC"] ;\n GT [label="GT"] ;\n CG [label="CG"] ;\n AC -> CG ;\n CG -> GC ;\n GC -> CG ;\n CG -> GT ;\n GT -> TC ;\n TC -> CG ;\n}\n'

runfile('C:/CS/bio/scratch.py', wdir='C:/CS/bio')

Traceback (most recent call last):

File "<ipython-input-20-1421c30afd69>", line 1, in <module>

runfile('C:/CS/bio/scratch.py', wdir='C:/CS/bio')

File "C:\CS\Anaconda\lib\site-packages\spyder\utils\site\sitecustomize.py", line 710, in runfile

execfile(filename, namespace)

File "C:\CS\Anaconda\lib\site-packages\spyder\utils\site\sitecustomize.py", line 101, in execfile

exec(compile(f.read(), filename, 'exec'), namespace)

File "C:/CS/bio/scratch.py", line 73, in <module>

f = sys.argv[1]

IndexError: list index out of range

runfile('C:/CS/bio/scratch.py', args='fragments.txt', wdir='C:/CS/bio')

TTACCACCAGGAATTACCAATTTACCAGGAGATTAGATTACCAGGATTACCCAGGACAATTACCAATTACCAGACCAGGATACCAGGACCAATTACCCAATTCCAATTACCAGGAACCAGGATACCAGAATTACCAGG

listsize= 6

length = 8

3

6

6

listsize= 134

length = 138

['AACCA', 'AATTA', 'AATTA', 'AATTA', 'AATTA', 'AATTA', 'AATTA', 'AATTC', 'AATTT', 'ACAAT', 'ACCAA', 'ACCAA', 'ACCAA', 'ACCAC', 'ACCAG', 'ACCAG', 'ACCAG', 'ACCAG', 'ACCAG', 'ACCAG', 'ACCAG', 'ACCAG', 'ACCAG', 'ACCAG', 'ACCCA', 'ACCCA', 'AGAAT', 'AGACC', 'AGATT', 'AGATT', 'AGGAA', 'AGGAA', 'AGGAC', 'AGGAC', 'AGGAG', 'AGGAT', 'AGGAT', 'AGGAT', 'ATACC', 'ATACC', 'ATTAC', 'ATTAC', 'ATTAC', 'ATTAC', 'ATTAC', 'ATTAC', 'ATTAC', 'ATTAC', 'ATTAG', 'ATTCC', 'ATTTA', 'CAATT', 'CAATT', 'CAATT', 'CAATT', 'CAATT', 'CAATT', 'CACCA', 'CAGAA', 'CAGAC', 'CAGGA', 'CAGGA', 'CAGGA', 'CAGGA', 'CAGGA', 'CAGGA', 'CAGGA', 'CAGGA', 'CCAAT', 'CCAAT', 'CCAAT', 'CCAAT', 'CCAAT', 'CCACC', 'CCAGA', 'CCAGA', 'CCAGG', 'CCAGG', 'CCAGG', 'CCAGG', 'CCAGG', 'CCAGG', 'CCAGG', 'CCAGG', 'CCAGG', 'CCCAA', 'CCCAG', 'GAACC', 'GAATT', 'GAATT', 'GACAA', 'GACCA', 'GACCA', 'GAGAT', 'GATAC', 'GATAC', 'GATTA', 'GATTA', 'GATTA', 'GGAAC', 'GGAAT', 'GGACA', 'GGACC', 'GGAGA', 'GGATA', 'GGATA', 'GGATT', 'TACCA', 'TACCA', 'TACCA', 'TACCA', 'TACCA', 'TACCA', 'TACCA', 'TACCA', 'TACCA', 'TACCA', 'TACCC', 'TACCC', 'TAGAT', 'TCCAA', 'TTACC', 'TTACC', 'TTACC', 'TTACC', 'TTACC', 'TTACC', 'TTACC', 'TTACC', 'TTACC', 'TTACC', 'TTAGA', 'TTCCA', 'TTTAC']

TTACCA

CCAGGA

ATTACCAATT

TACCAGGA

GATTA

GATTA

CCAGG

ATTAC

CCAGGA

CAATTACC

AATTACCAG

ACCAGGA

TACCAGG

ACCAATTAC

CCAATT

CCAATTA

CCAGGA

ACCAGGA

TACCAG

AATTACCAGG

frags

Traceback (most recent call last):

File "<ipython-input-22-2458ab5c9d00>", line 1, in <module>

frags

NameError: name 'frags' is not defined

runfile('C:/CS/bio/scratch.py', args='fragments.txt', wdir='C:/CS/bio')

TTACCACCAGGAATTACCAATTTACCAGGAGATTAGATTACCAGGATTACCCAGGACAATTACCAATTACCAGACCAGGATACCAGGACCAATTACCCAATTCCAATTACCAGGAACCAGGATACCAGAATTACCAGG

listsize= 6

length = 8

3

6

6

listsize= 134

length = 138

['AACCA', 'AATTA', 'AATTA', 'AATTA', 'AATTA', 'AATTA', 'AATTA', 'AATTC', 'AATTT', 'ACAAT', 'ACCAA', 'ACCAA', 'ACCAA', 'ACCAC', 'ACCAG', 'ACCAG', 'ACCAG', 'ACCAG', 'ACCAG', 'ACCAG', 'ACCAG', 'ACCAG', 'ACCAG', 'ACCAG', 'ACCCA', 'ACCCA', 'AGAAT', 'AGACC', 'AGATT', 'AGATT', 'AGGAA', 'AGGAA', 'AGGAC', 'AGGAC', 'AGGAG', 'AGGAT', 'AGGAT', 'AGGAT', 'ATACC', 'ATACC', 'ATTAC', 'ATTAC', 'ATTAC', 'ATTAC', 'ATTAC', 'ATTAC', 'ATTAC', 'ATTAC', 'ATTAG', 'ATTCC', 'ATTTA', 'CAATT', 'CAATT', 'CAATT', 'CAATT', 'CAATT', 'CAATT', 'CACCA', 'CAGAA', 'CAGAC', 'CAGGA', 'CAGGA', 'CAGGA', 'CAGGA', 'CAGGA', 'CAGGA', 'CAGGA', 'CAGGA', 'CCAAT', 'CCAAT', 'CCAAT', 'CCAAT', 'CCAAT', 'CCACC', 'CCAGA', 'CCAGA', 'CCAGG', 'CCAGG', 'CCAGG', 'CCAGG', 'CCAGG', 'CCAGG', 'CCAGG', 'CCAGG', 'CCAGG', 'CCCAA', 'CCCAG', 'GAACC', 'GAATT', 'GAATT', 'GACAA', 'GACCA', 'GACCA', 'GAGAT', 'GATAC', 'GATAC', 'GATTA', 'GATTA', 'GATTA', 'GGAAC', 'GGAAT', 'GGACA', 'GGACC', 'GGAGA', 'GGATA', 'GGATA', 'GGATT', 'TACCA', 'TACCA', 'TACCA', 'TACCA', 'TACCA', 'TACCA', 'TACCA', 'TACCA', 'TACCA', 'TACCA', 'TACCC', 'TACCC', 'TAGAT', 'TCCAA', 'TTACC', 'TTACC', 'TTACC', 'TTACC', 'TTACC', 'TTACC', 'TTACC', 'TTACC', 'TTACC', 'TTACC', 'TTAGA', 'TTCCA', 'TTTAC']

Traceback (most recent call last):

File "<ipython-input-23-9167c36bdf34>", line 1, in <module>

runfile('C:/CS/bio/scratch.py', args='fragments.txt', wdir='C:/CS/bio')

File "C:\CS\Anaconda\lib\site-packages\spyder\utils\site\sitecustomize.py", line 710, in runfile

execfile(filename, namespace)

File "C:\CS\Anaconda\lib\site-packages\spyder\utils\site\sitecustomize.py", line 101, in execfile

exec(compile(f.read(), filename, 'exec'), namespace)

File "C:/CS/bio/scratch.py", line 83, in <module>

print(myconcat(fileToList(f)))

File "C:/CS/bio/scratch.py", line 29, in myconcat

file = open(frags, 'r')

OSError: [Errno 22] Invalid argument: 'TTACCA\nCCAGGA\nATTACCAATT\nTACCAGGA\nGATTA\nGATTA\nCCAGG\nATTAC\nCCAGGA\nCAATTACC\nAATTACCAG\nACCAGGA\nTACCAGG\nACCAATTAC\nCCAATT\nCCAATTA\nCCAGGA\nACCAGGA\nTACCAG\nAATTACCAGG'

runfile('C:/CS/bio/scratch.py', args='fragments.txt', wdir='C:/CS/bio')

TTACCACCAGGAATTACCAATTTACCAGGAGATTAGATTACCAGGATTACCCAGGACAATTACCAATTACCAGACCAGGATACCAGGACCAATTACCCAATTCCAATTACCAGGAACCAGGATACCAGAATTACCAGG

listsize= 6

length = 8

3

6

6

listsize= 134

length = 138

['AACCA', 'AATTA', 'AATTA', 'AATTA', 'AATTA', 'AATTA', 'AATTA', 'AATTC', 'AATTT', 'ACAAT', 'ACCAA', 'ACCAA', 'ACCAA', 'ACCAC', 'ACCAG', 'ACCAG', 'ACCAG', 'ACCAG', 'ACCAG', 'ACCAG', 'ACCAG', 'ACCAG', 'ACCAG', 'ACCAG', 'ACCCA', 'ACCCA', 'AGAAT', 'AGACC', 'AGATT', 'AGATT', 'AGGAA', 'AGGAA', 'AGGAC', 'AGGAC', 'AGGAG', 'AGGAT', 'AGGAT', 'AGGAT', 'ATACC', 'ATACC', 'ATTAC', 'ATTAC', 'ATTAC', 'ATTAC', 'ATTAC', 'ATTAC', 'ATTAC', 'ATTAC', 'ATTAG', 'ATTCC', 'ATTTA', 'CAATT', 'CAATT', 'CAATT', 'CAATT', 'CAATT', 'CAATT', 'CACCA', 'CAGAA', 'CAGAC', 'CAGGA', 'CAGGA', 'CAGGA', 'CAGGA', 'CAGGA', 'CAGGA', 'CAGGA', 'CAGGA', 'CCAAT', 'CCAAT', 'CCAAT', 'CCAAT', 'CCAAT', 'CCACC', 'CCAGA', 'CCAGA', 'CCAGG', 'CCAGG', 'CCAGG', 'CCAGG', 'CCAGG', 'CCAGG', 'CCAGG', 'CCAGG', 'CCAGG', 'CCCAA', 'CCCAG', 'GAACC', 'GAATT', 'GAATT', 'GACAA', 'GACCA', 'GACCA', 'GAGAT', 'GATAC', 'GATAC', 'GATTA', 'GATTA', 'GATTA', 'GGAAC', 'GGAAT', 'GGACA', 'GGACC', 'GGAGA', 'GGATA', 'GGATA', 'GGATT', 'TACCA', 'TACCA', 'TACCA', 'TACCA', 'TACCA', 'TACCA', 'TACCA', 'TACCA', 'TACCA', 'TACCA', 'TACCC', 'TACCC', 'TAGAT', 'TCCAA', 'TTACC', 'TTACC', 'TTACC', 'TTACC', 'TTACC', 'TTACC', 'TTACC', 'TTACC', 'TTACC', 'TTACC', 'TTAGA', 'TTCCA', 'TTTAC']

TTACCACCAGGAATTACCAATTTACCAGGAGATTAGATTACCAGGATTACCCAGGACAATTACCAATTACCAGACCAGGATACCAGGACCAATTACCCAATTCCAATTACCAGGAACCAGGATACCAGAATTACCAGG

runfile('C:/CS/bio/scratch.py', args='fragments.txt', wdir='C:/CS/bio')

TTACCACCAGGAATTACCAATTTACCAGGAGATTAGATTACCAGGATTACCCAGGACAATTACCAATTACCAGACCAGGATACCAGGACCAATTACCCAATTCCAATTACCAGGAACCAGGATACCAGAATTACCAGG

listsize= 6

length = 8

3

6

6

listsize= 136

length = 138

['AAC', 'AAT', 'AAT', 'AAT', 'AAT', 'AAT', 'AAT', 'AAT', 'AAT', 'ACA', 'ACC', 'ACC', 'ACC', 'ACC', 'ACC', 'ACC', 'ACC', 'ACC', 'ACC', 'ACC', 'ACC', 'ACC', 'ACC', 'ACC', 'ACC', 'ACC', 'AGA', 'AGA', 'AGA', 'AGA', 'AGG', 'AGG', 'AGG', 'AGG', 'AGG', 'AGG', 'AGG', 'AGG', 'AGG', 'ATA', 'ATA', 'ATT', 'ATT', 'ATT', 'ATT', 'ATT', 'ATT', 'ATT', 'ATT', 'ATT', 'ATT', 'ATT', 'CAA', 'CAA', 'CAA', 'CAA', 'CAA', 'CAA', 'CAC', 'CAG', 'CAG', 'CAG', 'CAG', 'CAG', 'CAG', 'CAG', 'CAG', 'CAG', 'CAG', 'CAG', 'CCA', 'CCA', 'CCA', 'CCA', 'CCA', 'CCA', 'CCA', 'CCA', 'CCA', 'CCA', 'CCA', 'CCA', 'CCA', 'CCA', 'CCA', 'CCA', 'CCA', 'CCC', 'CCC', 'GAA', 'GAA', 'GAA', 'GAC', 'GAC', 'GAC', 'GAG', 'GAT', 'GAT', 'GAT', 'GAT', 'GAT', 'GGA', 'GGA', 'GGA', 'GGA', 'GGA', 'GGA', 'GGA', 'GGA', 'TAC', 'TAC', 'TAC', 'TAC', 'TAC', 'TAC', 'TAC', 'TAC', 'TAC', 'TAC', 'TAC', 'TAC', 'TAG', 'TCC', 'TTA', 'TTA', 'TTA', 'TTA', 'TTA', 'TTA', 'TTA', 'TTA', 'TTA', 'TTA', 'TTA', 'TTC', 'TTT']

TTACCACCAGGAATTACCAATTTACCAGGAGATTAGATTACCAGGATTACCCAGGACAATTACCAATTACCAGACCAGGATACCAGGACCAATTACCCAATTCCAATTACCAGGAACCAGGATACCAGAATTACCAGG

runfile('C:/CS/bio/debruijn.py', wdir='C:/CS/bio')

debrujn(s2)

Traceback (most recent call last):

File "<ipython-input-27-55e2d8dde0fa>", line 1, in <module>

debrujn(s2)

NameError: name 'debrujn' is not defined

debruijn(s2)

Traceback (most recent call last):

File "<ipython-input-28-7f8ffe4c50cc>", line 1, in <module>

debruijn(s2)

TypeError: debruijn() missing 1 required positional argument: 'k'

debruijn(s2,3)

Out[29]:

({'AA', 'AC', 'AG', 'AT', 'CA', 'CC', 'GA', 'GG', 'TA', 'TC', 'TT'},

[('TT', 'TA'),

('TA', 'AC'),

('AC', 'CC'),

('CC', 'CA'),

('CA', 'AC'),

('AC', 'CC'),

('CC', 'CA'),

('CA', 'AG'),

('AG', 'GG'),

('GG', 'GA'),

('GA', 'AA'),

('AA', 'AT'),

('AT', 'TT'),

('TT', 'TA'),

('TA', 'AC'),

('AC', 'CC'),

('CC', 'CA'),

('CA', 'AA'),

('AA', 'AT'),

('AT', 'TT'),

('TT', 'TT'),

('TT', 'TA'),

('TA', 'AC'),

('AC', 'CC'),

('CC', 'CA'),

('CA', 'AG'),

('AG', 'GG'),

('GG', 'GA'),

('GA', 'AG'),

('AG', 'GA'),

('GA', 'AT'),

('AT', 'TT'),

('TT', 'TA'),

('TA', 'AG'),

('AG', 'GA'),

('GA', 'AT'),

('AT', 'TT'),

('TT', 'TA'),

('TA', 'AC'),

('AC', 'CC'),

('CC', 'CA'),

('CA', 'AG'),

('AG', 'GG'),

('GG', 'GA'),

('GA', 'AT'),

('AT', 'TT'),

('TT', 'TA'),

('TA', 'AC'),

('AC', 'CC'),

('CC', 'CC'),

('CC', 'CA'),

('CA', 'AG'),

('AG', 'GG'),

('GG', 'GA'),

('GA', 'AC'),

('AC', 'CA'),

('CA', 'AA'),

('AA', 'AT'),

('AT', 'TT'),

('TT', 'TA'),

('TA', 'AC'),

('AC', 'CC'),

('CC', 'CA'),

('CA', 'AA'),

('AA', 'AT'),

('AT', 'TT'),

('TT', 'TA'),

('TA', 'AC'),

('AC', 'CC'),

('CC', 'CA'),

('CA', 'AG'),

('AG', 'GA'),

('GA', 'AC'),

('AC', 'CC'),

('CC', 'CA'),

('CA', 'AG'),

('AG', 'GG'),

('GG', 'GA'),

('GA', 'AT'),

('AT', 'TA'),

('TA', 'AC'),

('AC', 'CC'),

('CC', 'CA'),

('CA', 'AG'),

('AG', 'GG'),

('GG', 'GA'),

('GA', 'AC'),

('AC', 'CC'),

('CC', 'CA'),

('CA', 'AA'),

('AA', 'AT'),

('AT', 'TT'),

('TT', 'TA'),

('TA', 'AC'),

('AC', 'CC'),

('CC', 'CC'),

('CC', 'CA'),

('CA', 'AA'),

('AA', 'AT'),

('AT', 'TT'),

('TT', 'TC'),

('TC', 'CC'),

('CC', 'CA'),

('CA', 'AA'),

('AA', 'AT'),

('AT', 'TT'),

('TT', 'TA'),

('TA', 'AC'),

('AC', 'CC'),

('CC', 'CA'),

('CA', 'AG'),

('AG', 'GG'),

('GG', 'GA'),

('GA', 'AA'),

('AA', 'AC'),

('AC', 'CC'),

('CC', 'CA'),

('CA', 'AG'),

('AG', 'GG'),

('GG', 'GA'),

('GA', 'AT'),

('AT', 'TA'),

('TA', 'AC'),

('AC', 'CC'),

('CC', 'CA'),

('CA', 'AG'),

('AG', 'GA'),

('GA', 'AA'),

('AA', 'AT'),

('AT', 'TT'),

('TT', 'TA'),

('TA', 'AC'),

('AC', 'CC'),

('CC', 'CA'),

('CA', 'AG'),

('AG', 'GG')])

visualize\_de\_bruijn(s2,3)

Out[30]: 'digraph "DeBruijn graph" {\n AT [label="AT"] ;\n CA [label="CA"] ;\n AC [label="AC"] ;\n CC [label="CC"] ;\n GG [label="GG"] ;\n GA [label="GA"] ;\n AA [label="AA"] ;\n TC [label="TC"] ;\n AG [label="AG"] ;\n TT [label="TT"] ;\n TA [label="TA"] ;\n TT -> TA ;\n TA -> AC ;\n AC -> CC ;\n CC -> CA ;\n CA -> AC ;\n AC -> CC ;\n CC -> CA ;\n CA -> AG ;\n AG -> GG ;\n GG -> GA ;\n GA -> AA ;\n AA -> AT ;\n AT -> TT ;\n TT -> TA ;\n TA -> AC ;\n AC -> CC ;\n CC -> CA ;\n CA -> AA ;\n AA -> AT ;\n AT -> TT ;\n TT -> TT ;\n TT -> TA ;\n TA -> AC ;\n AC -> CC ;\n CC -> CA ;\n CA -> AG ;\n AG -> GG ;\n GG -> GA ;\n GA -> AG ;\n AG -> GA ;\n GA -> AT ;\n AT -> TT ;\n TT -> TA ;\n TA -> AG ;\n AG -> GA ;\n GA -> AT ;\n AT -> TT ;\n TT -> TA ;\n TA -> AC ;\n AC -> CC ;\n CC -> CA ;\n CA -> AG ;\n AG -> GG ;\n GG -> GA ;\n GA -> AT ;\n AT -> TT ;\n TT -> TA ;\n TA -> AC ;\n AC -> CC ;\n CC -> CC ;\n CC -> CA ;\n CA -> AG ;\n AG -> GG ;\n GG -> GA ;\n GA -> AC ;\n AC -> CA ;\n CA -> AA ;\n AA -> AT ;\n AT -> TT ;\n TT -> TA ;\n TA -> AC ;\n AC -> CC ;\n CC -> CA ;\n CA -> AA ;\n AA -> AT ;\n AT -> TT ;\n TT -> TA ;\n TA -> AC ;\n AC -> CC ;\n CC -> CA ;\n CA -> AG ;\n AG -> GA ;\n GA -> AC ;\n AC -> CC ;\n CC -> CA ;\n CA -> AG ;\n AG -> GG ;\n GG -> GA ;\n GA -> AT ;\n AT -> TA ;\n TA -> AC ;\n AC -> CC ;\n CC -> CA ;\n CA -> AG ;\n AG -> GG ;\n GG -> GA ;\n GA -> AC ;\n AC -> CC ;\n CC -> CA ;\n CA -> AA ;\n AA -> AT ;\n AT -> TT ;\n TT -> TA ;\n TA -> AC ;\n AC -> CC ;\n CC -> CC ;\n CC -> CA ;\n CA -> AA ;\n AA -> AT ;\n AT -> TT ;\n TT -> TC ;\n TC -> CC ;\n CC -> CA ;\n CA -> AA ;\n AA -> AT ;\n AT -> TT ;\n TT -> TA ;\n TA -> AC ;\n AC -> CC ;\n CC -> CA ;\n CA -> AG ;\n AG -> GG ;\n GG -> GA ;\n GA -> AA ;\n AA -> AC ;\n AC -> CC ;\n CC -> CA ;\n CA -> AG ;\n AG -> GG ;\n GG -> GA ;\n GA -> AT ;\n AT -> TA ;\n TA -> AC ;\n AC -> CC ;\n CC -> CA ;\n CA -> AG ;\n AG -> GA ;\n GA -> AA ;\n AA -> AT ;\n AT -> TT ;\n TT -> TA ;\n TA -> AC ;\n AC -> CC ;\n CC -> CA ;\n CA -> AG ;\n AG -> GG ;\n}\n'

visualize\_de\_bruijn(s2,3)

Out[31]: 'digraph "DeBruijn graph" {\n AT [label="AT"] ;\n CA [label="CA"] ;\n AC [label="AC"] ;\n CC [label="CC"] ;\n GG [label="GG"] ;\n GA [label="GA"] ;\n AA [label="AA"] ;\n TC [label="TC"] ;\n AG [label="AG"] ;\n TT [label="TT"] ;\n TA [label="TA"] ;\n TT -> TA ;\n TA -> AC ;\n AC -> CC ;\n CC -> CA ;\n CA -> AC ;\n AC -> CC ;\n CC -> CA ;\n CA -> AG ;\n AG -> GG ;\n GG -> GA ;\n GA -> AA ;\n AA -> AT ;\n AT -> TT ;\n TT -> TA ;\n TA -> AC ;\n AC -> CC ;\n CC -> CA ;\n CA -> AA ;\n AA -> AT ;\n AT -> TT ;\n TT -> TT ;\n TT -> TA ;\n TA -> AC ;\n AC -> CC ;\n CC -> CA ;\n CA -> AG ;\n AG -> GG ;\n GG -> GA ;\n GA -> AG ;\n AG -> GA ;\n GA -> AT ;\n AT -> TT ;\n TT -> TA ;\n TA -> AG ;\n AG -> GA ;\n GA -> AT ;\n AT -> TT ;\n TT -> TA ;\n TA -> AC ;\n AC -> CC ;\n CC -> CA ;\n CA -> AG ;\n AG -> GG ;\n GG -> GA ;\n GA -> AT ;\n AT -> TT ;\n TT -> TA ;\n TA -> AC ;\n AC -> CC ;\n CC -> CC ;\n CC -> CA ;\n CA -> AG ;\n AG -> GG ;\n GG -> GA ;\n GA -> AC ;\n AC -> CA ;\n CA -> AA ;\n AA -> AT ;\n AT -> TT ;\n TT -> TA ;\n TA -> AC ;\n AC -> CC ;\n CC -> CA ;\n CA -> AA ;\n AA -> AT ;\n AT -> TT ;\n TT -> TA ;\n TA -> AC ;\n AC -> CC ;\n CC -> CA ;\n CA -> AG ;\n AG -> GA ;\n GA -> AC ;\n AC -> CC ;\n CC -> CA ;\n CA -> AG ;\n AG -> GG ;\n GG -> GA ;\n GA -> AT ;\n AT -> TA ;\n TA -> AC ;\n AC -> CC ;\n CC -> CA ;\n CA -> AG ;\n AG -> GG ;\n GG -> GA ;\n GA -> AC ;\n AC -> CC ;\n CC -> CA ;\n CA -> AA ;\n AA -> AT ;\n AT -> TT ;\n TT -> TA ;\n TA -> AC ;\n AC -> CC ;\n CC -> CC ;\n CC -> CA ;\n CA -> AA ;\n AA -> AT ;\n AT -> TT ;\n TT -> TC ;\n TC -> CC ;\n CC -> CA ;\n CA -> AA ;\n AA -> AT ;\n AT -> TT ;\n TT -> TA ;\n TA -> AC ;\n AC -> CC ;\n CC -> CA ;\n CA -> AG ;\n AG -> GG ;\n GG -> GA ;\n GA -> AA ;\n AA -> AC ;\n AC -> CC ;\n CC -> CA ;\n CA -> AG ;\n AG -> GG ;\n GG -> GA ;\n GA -> AT ;\n AT -> TA ;\n TA -> AC ;\n AC -> CC ;\n CC -> CA ;\n CA -> AG ;\n AG -> GA ;\n GA -> AA ;\n AA -> AT ;\n AT -> TT ;\n TT -> TA ;\n TA -> AC ;\n AC -> CC ;\n CC -> CA ;\n CA -> AG ;\n AG -> GG ;\n}\n'

runfile('C:/CS/bio/debruijn.py', wdir='C:/CS/bio')

runfile('C:/CS/bio/debruijn.py', wdir='C:/CS/bio')

visualize\_de\_bruijn(s2,3)

Out[34]: 'digraph "DeBruijn graph" { AT [label="AT"] ; CA [label="CA"] ; AC [label="AC"] ; CC [label="CC"] ; GG [label="GG"] ; GA [label="GA"] ; AA [label="AA"] ; TC [label="TC"] ; AG [label="AG"] ; TT [label="TT"] ; TA [label="TA"] ; TT -> TA ; TA -> AC ; AC -> CC ; CC -> CA ; CA -> AC ; AC -> CC ; CC -> CA ; CA -> AG ; AG -> GG ; GG -> GA ; GA -> AA ; AA -> AT ; AT -> TT ; TT -> TA ; TA -> AC ; AC -> CC ; CC -> CA ; CA -> AA ; AA -> AT ; AT -> TT ; TT -> TT ; TT -> TA ; TA -> AC ; AC -> CC ; CC -> CA ; CA -> AG ; AG -> GG ; GG -> GA ; GA -> AG ; AG -> GA ; GA -> AT ; AT -> TT ; TT -> TA ; TA -> AG ; AG -> GA ; GA -> AT ; AT -> TT ; TT -> TA ; TA -> AC ; AC -> CC ; CC -> CA ; CA -> AG ; AG -> GG ; GG -> GA ; GA -> AT ; AT -> TT ; TT -> TA ; TA -> AC ; AC -> CC ; CC -> CC ; CC -> CA ; CA -> AG ; AG -> GG ; GG -> GA ; GA -> AC ; AC -> CA ; CA -> AA ; AA -> AT ; AT -> TT ; TT -> TA ; TA -> AC ; AC -> CC ; CC -> CA ; CA -> AA ; AA -> AT ; AT -> TT ; TT -> TA ; TA -> AC ; AC -> CC ; CC -> CA ; CA -> AG ; AG -> GA ; GA -> AC ; AC -> CC ; CC -> CA ; CA -> AG ; AG -> GG ; GG -> GA ; GA -> AT ; AT -> TA ; TA -> AC ; AC -> CC ; CC -> CA ; CA -> AG ; AG -> GG ; GG -> GA ; GA -> AC ; AC -> CC ; CC -> CA ; CA -> AA ; AA -> AT ; AT -> TT ; TT -> TA ; TA -> AC ; AC -> CC ; CC -> CC ; CC -> CA ; CA -> AA ; AA -> AT ; AT -> TT ; TT -> TC ; TC -> CC ; CC -> CA ; CA -> AA ; AA -> AT ; AT -> TT ; TT -> TA ; TA -> AC ; AC -> CC ; CC -> CA ; CA -> AG ; AG -> GG ; GG -> GA ; GA -> AA ; AA -> AC ; AC -> CC ; CC -> CA ; CA -> AG ; AG -> GG ; GG -> GA ; GA -> AT ; AT -> TA ; TA -> AC ; AC -> CC ; CC -> CA ; CA -> AG ; AG -> GA ; GA -> AA ; AA -> AT ; AT -> TT ; TT -> TA ; TA -> AC ; AC -> CC ; CC -> CA ; CA -> AG ; AG -> GG ;}'

runfile('C:/CS/bio/make\_fragments.py', wdir='C:/CS/bio')

usage: <genome> <number\_of\_reads> <min\_read\_length> max\_read\_length>

make\_fragments.py TAATGCCATGGGATGTT 20 5 10

File "<ipython-input-36-2c3a4bee878f>", line 1

make\_fragments.py TAATGCCATGGGATGTT 20 5 10

^

SyntaxError: invalid syntax

make\_fragments TAATGCCATGGGATGTT 20 5 10

File "<ipython-input-37-c2393245df2b>", line 1

make\_fragments TAATGCCATGGGATGTT 20 5 10

^

SyntaxError: invalid syntax

make\_fragments 'TAATGCCATGGGATGTT' 20 5 10

File "<ipython-input-38-bcb01c632589>", line 1

make\_fragments 'TAATGCCATGGGATGTT' 20 5 10

^

SyntaxError: invalid syntax

make\_fragments.py 'TAATGCCATGGGATGTT' 20 5 10

File "<ipython-input-39-dd66f348874d>", line 1

make\_fragments.py 'TAATGCCATGGGATGTT' 20 5 10

^

SyntaxError: invalid syntax

make\_fragments.py

Traceback (most recent call last):

File "<ipython-input-40-9bef5b89d7f0>", line 1, in <module>

make\_fragments.py

NameError: name 'make\_fragments' is not defined

runfile('C:/CS/bio/make\_fragments.py', wdir='C:/CS/bio')

usage: <genome> <number\_of\_reads> <min\_read\_length> max\_read\_length>

runfile('C:/CS/bio/make\_fragments.py', args='TAATGCCATGGGATGTT 15 5 10', wdir='C:/CS/bio')

AATGCCA

TAATGCCATG

GATGTT

CATGGGAT

GGATGTT

TAATGCC

GGGATGTT

TGGGATGTT

ATGGGAT

AATGCCAT

GCCATG

TGCCATGG

GATGTT

TGCCATGGG

GGGATGT

runfile('C:/CS/bio/scratch.py', args='TAATGCCATGGGATGTT.txt', wdir='C:/CS/bio')

AATGCCATAATGCCATGGATGTTCATGGGATGGATGTTTAATGCCGGGATGTTTGGGATGTTATGGGATAATGCCATGCCATGTGCCATGGGATGTTTGCCATGGGGGGATGT

listsize= 6

length = 8

3

6

6

listsize= 111

length = 113

['AAT', 'AAT', 'AAT', 'AAT', 'ATA', 'ATA', 'ATG', 'ATG', 'ATG', 'ATG', 'ATG', 'ATG', 'ATG', 'ATG', 'ATG', 'ATG', 'ATG', 'ATG', 'ATG', 'ATG', 'ATG', 'ATG', 'ATG', 'ATG', 'CAT', 'CAT', 'CAT', 'CAT', 'CAT', 'CAT', 'CAT', 'CCA', 'CCA', 'CCA', 'CCA', 'CCA', 'CCA', 'CCG', 'CGG', 'GAT', 'GAT', 'GAT', 'GAT', 'GAT', 'GAT', 'GAT', 'GAT', 'GCC', 'GCC', 'GCC', 'GCC', 'GCC', 'GCC', 'GCC', 'GGA', 'GGA', 'GGA', 'GGA', 'GGA', 'GGA', 'GGA', 'GGA', 'GGG', 'GGG', 'GGG', 'GGG', 'GGG', 'GGG', 'GGG', 'GGG', 'GGG', 'GTG', 'GTT', 'GTT', 'GTT', 'GTT', 'GTT', 'TAA', 'TAA', 'TAA', 'TAT', 'TCA', 'TGC', 'TGC', 'TGC', 'TGC', 'TGC', 'TGC', 'TGC', 'TGG', 'TGG', 'TGG', 'TGG', 'TGG', 'TGG', 'TGG', 'TGT', 'TGT', 'TGT', 'TGT', 'TGT', 'TGT', 'TGT', 'TTA', 'TTA', 'TTC', 'TTG', 'TTG', 'TTT', 'TTT', 'TTT']

AATGCCATAATGCCATGGATGTTCATGGGATGGATGTTTAATGCCGGGATGTTTGGGATGTTATGGGATAATGCCATGCCATGTGCCATGGGATGTTTGCCATGGGGGGATGT

s2

Out[44]: 'AATGCCATAATGCCATGGATGTTCATGGGATGGATGTTTAATGCCGGGATGTTTGGGATGTTATGGGATAATGCCATGCCATGTGCCATGGGATGTTTGCCATGGGGGGATGT'

visualize\_de\_bruijn(Out[44],3)

Out[45]: 'digraph "DeBruijn graph" { AT [label="AT"] ; CA [label="CA"] ; CC [label="CC"] ; GC [label="GC"] ; GG [label="GG"] ; AA [label="AA"] ; GA [label="GA"] ; TC [label="TC"] ; GT [label="GT"] ; TG [label="TG"] ; CG [label="CG"] ; TT [label="TT"] ; TA [label="TA"] ; AA -> AT ; AT -> TG ; TG -> GC ; GC -> CC ; CC -> CA ; CA -> AT ; AT -> TA ; TA -> AA ; AA -> AT ; AT -> TG ; TG -> GC ; GC -> CC ; CC -> CA ; CA -> AT ; AT -> TG ; TG -> GG ; GG -> GA ; GA -> AT ; AT -> TG ; TG -> GT ; GT -> TT ; TT -> TC ; TC -> CA ; CA -> AT ; AT -> TG ; TG -> GG ; GG -> GG ; GG -> GA ; GA -> AT ; AT -> TG ; TG -> GG ; GG -> GA ; GA -> AT ; AT -> TG ; TG -> GT ; GT -> TT ; TT -> TT ; TT -> TA ; TA -> AA ; AA -> AT ; AT -> TG ; TG -> GC ; GC -> CC ; CC -> CG ; CG -> GG ; GG -> GG ; GG -> GA ; GA -> AT ; AT -> TG ; TG -> GT ; GT -> TT ; TT -> TT ; TT -> TG ; TG -> GG ; GG -> GG ; GG -> GA ; GA -> AT ; AT -> TG ; TG -> GT ; GT -> TT ; TT -> TA ; TA -> AT ; AT -> TG ; TG -> GG ; GG -> GG ; GG -> GA ; GA -> AT ; AT -> TA ; TA -> AA ; AA -> AT ; AT -> TG ; TG -> GC ; GC -> CC ; CC -> CA ; CA -> AT ; AT -> TG ; TG -> GC ; GC -> CC ; CC -> CA ; CA -> AT ; AT -> TG ; TG -> GT ; GT -> TG ; TG -> GC ; GC -> CC ; CC -> CA ; CA -> AT ; AT -> TG ; TG -> GG ; GG -> GG ; GG -> GA ; GA -> AT ; AT -> TG ; TG -> GT ; GT -> TT ; TT -> TT ; TT -> TG ; TG -> GC ; GC -> CC ; CC -> CA ; CA -> AT ; AT -> TG ; TG -> GG ; GG -> GG ; GG -> GG ; GG -> GG ; GG -> GG ; GG -> GA ; GA -> AT ; AT -> TG ; TG -> GT ;}'