CSCE 638: Natural Language Processing

Assignment 3 Report

Submitted By – Sameer Kumar Behera UIN: 526004296

System Requirements:

- Python3 must be installed (Python2 Not Recommended)
- Installation Link: https://www.python.org/downloads/release/python-370/

Compile and Run Method:

1. Open a Terminal
2. Go to the Project Folder, i.e. PA3-638
3. On the Terminal, Enter the Following Command:

   - **python3 cky.py grammar_rules.txt sents.txt**

Result and Analysis:

The CKY program was run with the given test case, i.e. {fish people fish tanks}. The implementation for CKY program outputs the correct results. Also, other different test cases, i.e. {people with fish tanks}, {people with fish rods}, etc. were added in the sents.txt file attached with the submission for more testing.

Clearly, CKY algorithm is a dynamic algorithm to efficiently find the most likely syntactic parse tree for a sentence.

Following results were obtained on running the program with above test cases:

PROCESSING SENTENCE: fish people fish tanks

SPAN: fish
P(V fish) = 0.6
P(N fish) = 0.2
P(S) = 0.006 (BackPointer = (VP))
P(VP) = 0.06 (BackPointer = (V))
P(NP) = 0.13999999999999999 (BackPointer = (N))

SPAN: people
P(V people) = 0.1
P(N people) = 0.5
P(S) = 0.0010000000000000002 (BackPointer = (VP))
P(VP) = 0.010000000000000002 (BackPointer = (V))
P(NP) = 0.35 (BackPointer = (N))

SPAN: fish
P(V fish) = 0.6
P(N fish) = 0.2
P(S) = 0.006 (BackPointer = (VP))
P(VP) = 0.06 (BackPointer = (V))
P(NP) = 0.13999999999999999 (BackPointer = (N))

SPAN: tanks
P(V tanks) = 0.3
P(N tanks) = 0.2
P(S) = 0.003 (BackPointer = (VP))
P(VP) = 0.03 (BackPointer = (V))
P(NP) = 0.13999999999999999 (BackPointer = (N))

SPAN: fish people
P(S) = 0.0105 (BackPointer = (VP))
P(VP) = 0.105 (BackPointer = (1, V, NP))
P(NP) = 0.0049 (BackPointer = (1, NP, NP))

SPAN: people fish
P(S) = 0.0189 (BackPointer = (2, NP, VP))
P(VP) = 0.006999999999999999 (BackPointer = (2, V, NP))
P(NP) = 0.0049 (BackPointer = (2, NP, NP))

SPAN: fish tanks
P(S) = 0.0042 (BackPointer = (VP))
P(VP) = 0.041999999999999996 (BackPointer = (3, V, NP))
P(NP) = 0.0019599999999999995 (BackPointer = (3, NP, NP))

SPAN: fish people fish
P(S) = 0.0008819999999999998 (BackPointer = (1, NP, VP))
P(VP) = 0.00147 (BackPointer = (1, V, NP))
P(NP) = 6.859999999999999e-05 (BackPointer = (1, NP, NP))

SPAN: people fish tanks
P(S) = 0.013229999999999999 (BackPointer = (2, NP, VP))
P(VP) = 9.799999999999998e-05 (BackPointer = (2, V, NP))
P(NP) = 6.859999999999999e-05 (BackPointer = (3, NP, NP))

SPAN: fish people fish tanks
P(S) = 0.00018521999999999996 (BackPointer = (2, NP, VP))
P(VP) = 2.0579999999999996e-05 (BackPointer = (1, V, NP))
P(NP) = 9.603999999999997e-07 (BackPointer = (1, NP, NP))

PROCESSING SENTENCE: people with fish tanks

SPAN: people
P(N people) = 0.5
P(V people) = 0.1
P(S) = 0.0010000000000000002 (BackPointer = (VP))
P(NP) = 0.35 (BackPointer = (N))
P(VP) = 0.010000000000000002 (BackPointer = (V))

SPAN: with
P(P with) = 1.0

SPAN: fish
P(N fish) = 0.2
P(V fish) = 0.6
P(S) = 0.006 (BackPointer = (VP))
P(NP) = 0.13999999999999999 (BackPointer = (N))
P(VP) = 0.06 (BackPointer = (V))

SPAN: tanks
P(N tanks) = 0.2
P(V tanks) = 0.3
P(S) = 0.003 (BackPointer = (VP))
P(NP) = 0.13999999999999999 (BackPointer = (N))
P(VP) = 0.03 (BackPointer = (V))

SPAN: people with

SPAN: with fish
P(PP) = 0.13999999999999999 (BackPointer = (2, P, NP))

SPAN: fish tanks
P(S) = 0.0042 (BackPointer = (VP))
P(NP) = 0.0019599999999999995 (BackPointer = (3, NP, NP))
P(VP) = 0.041999999999999996 (BackPointer = (3, V, NP))

SPAN: people with fish
P(@VP_V) = 0.048999999999999995 (BackPointer = (1, NP, PP))
P(S) = 0.00014000000000000001 (BackPointer = (VP))
P(NP) = 0.0098 (BackPointer = (1, NP, PP))
P(VP) = 0.0014 (BackPointer = (1, V, PP))

SPAN: with fish tanks
P(PP) = 0.0019599999999999995 (BackPointer = (2, P, NP))

SPAN: people with fish tanks
P(@VP_V) = 0.0006859999999999998 (BackPointer = (1, NP, PP))
P(S) = 0.0002646 (BackPointer = (3, NP, VP))
P(NP) = 0.00013719999999999997 (BackPointer = (3, NP, NP))
P(VP) = 1.96e-05 (BackPointer = (1, V, PP))

PROCESSING SENTENCE: people with fish rods

SPAN: people
P(N people) = 0.5
P(V people) = 0.1
P(S) = 0.0010000000000000002 (BackPointer = (VP))
P(NP) = 0.35 (BackPointer = (N))
P(VP) = 0.010000000000000002 (BackPointer = (V))

SPAN: with

P(P with) = 1.0

SPAN: fish
P(N fish) = 0.2
P(V fish) = 0.6
P(S) = 0.006 (BackPointer = (VP))
P(NP) = 0.13999999999999999 (BackPointer = (N))
P(VP) = 0.06 (BackPointer = (V))

SPAN: rods
P(N rods) = 0.1
P(NP) = 0.06999999999999999 (BackPointer = (N))

SPAN: people with

SPAN: with fish
P(PP) = 0.13999999999999999 (BackPointer = (2, P, NP))

SPAN: fish rods
P(S) = 0.0021 (BackPointer = (VP))
P(NP) = 0.0009799999999999998 (BackPointer = (3, NP, NP))
P(VP) = 0.020999999999999998 (BackPointer = (3, V, NP))

SPAN: people with fish
P(@VP_V) = 0.048999999999999995 (BackPointer = (1, NP, PP))
P(S) = 0.00014000000000000001 (BackPointer = (VP))
P(NP) = 0.0098 (BackPointer = (1, NP, PP))
P(VP) = 0.0014 (BackPointer = (1, V, PP))

SPAN: with fish rods
P(PP) = 0.0009799999999999998 (BackPointer = (2, P, NP))

SPAN: people with fish rods
P(@VP_V) = 0.0003429999999999999 (BackPointer = (1, NP, PP))
P(S) = 9.8e-07 (BackPointer = (VP))
P(NP) = 6.859999999999999e-05 (BackPointer = (3, NP, NP))
P(VP) = 9.8e-06 (BackPointer = (1, V, PP))

Any Known Bugs or Limitations:

No such bugs/limitation were found were found in the code.

References Referred To:

- https://en.wikipedia.org/wiki/CYK_algorithm
- https://www.xarg.org/tools/cyk-algorithm
- http://courses.washington.edu/ling571/ling571_WIN2016/slides/ling571_class3_cnf_cky_flat.pdf