

Features Chosen: pB , pI

pB = probability that current token is the beginning of a noun group {0,1}

pI = probability that current token is inside of a noun group {0,1}

priorTag, currentTag, nextTag determine the truth value of the features.

The term “noun-group” includes:

NN, NNP, NNS, DT, JJ, PRP

#if a noun-group token is preceded by whitespace or a non noun-group token

if (((priorTag == "") or (priorTag != "NN" and priorTag != "NNP" and priorTag != "NNS" and priorTag != "DT" and priorTag != "JJ" and priorTag != "PRP")) and (currentTag == "NN" or currentTag == "NNP" or currentTag == "NNS" or currentTag == "JJ" or currentTag == "DT" or currentTag == "PRP"))):

#if a determiner or adjective

if (currentTag == "DT" or currentTag == "JJ"):

#precedes a non noun-group token

if (nextTag == "JJ" or nextTag == "NN" or nextTag == "NNP" or nextTag == "NNS"):

return "pB=1 pI=0 B-NP"

else:

return "pB=0 pI=0 O"

else:

return "pB=1 pI=0 B-NP"

#else if a noun-group token is preceded by a noun-group token

elif (priorTag == "NN" or priorTag == "NNP" or priorTag == "NNS" or priorTag == "JJ" or priorTag == "DT" or priorTag == "PRP") and (currentTag == "NN" or currentTag == "NNP" or currentTag == "NNS" or currentTag == "JJ" or currentTag == "DT" or currentTag == "PRP"):

return "pB=0 pI=1 I-NP"

else:

return "pB=0 pI=0 O"