

Richard Ngo  
Q2

FORMAT: ['SENTENCE', <SENTENCE NUMBER>, <PAIR>, <CONFIDENCE LEVEL>, 'CASE 1', <CASE 1 COUNT>, <LIST OF SYNSETS FOR CASE 1>, 'CASE 2', <CASE 2 COUNT>, <LIST OF SYNSETS FOR CASE 2>, 'CASE 3', <CASE 3 COUNT>, <LIST OF SYNSETS FOR CASE 3>, 'CASE 4', <CASE 4 COUNT>, <LIST OF SYNSETS FOR CASE 4>, 'RULE', <HEARST RULE FOR CURRENT PAIR> ]

Due to memory issues, program ran for nyt\_big (0:800000)

For each case, Rule starts from 1 to 5

#### CASE 1

['SENTENCE', 127875, ['fracture', 'injury'], 'LOW', 'CASE 1:', 1, [Synset('fracture.n.01')], 'CASE 2:', 0, [], 'CASE 3:', 0, 'CASE 4:', 0, [], 'RULE:', 1]

['SENTENCE', 210015, ['farming', 'occupation'], 'LOW', 'CASE 1:', 1, [Synset('farming.n.02')], 'CASE 2:', 0, [], 'CASE 3:', 0, 'CASE 4:', 0, [], 'RULE:', 1]

['SENTENCE', 278002, ['spring', 'season'], 'LOW', 'CASE 1:', 1, [Synset('spring.n.01')], 'CASE 2:', 0, [], 'CASE 3:', 0, 'CASE 4:', 0, [], 'RULE:', 1]

['SENTENCE', 211582, ['Daily', 'paper'], 'LOW', 'CASE 1:', 1, [Synset('daily.n.01')], 'CASE 2:', 0, [], 'CASE 3:', 0, 'CASE 4:', 0, [], 'RULE:', 2]

['SENTENCE', 2526, ['debt', 'obligation'], 'LOW', 'CASE 1:', 1, [Synset('debt.n.03')], 'CASE 2:', 0, [], 'CASE 3:', 0, 'CASE 4:', 0, [], 'RULE:', 3]

['SENTENCE', 15093, ['lyric', 'writing'], 'MEDIUM', 'CASE 1:', 1, [Synset('lyric.v.01')], 'CASE 2:', 0, [], 'CASE 3:', 0, 'CASE 4:', 0, [], 'RULE:', 3]

['SENTENCE', 20358, ['association', 'group'], 'MEDIUM', 'CASE 1:', 1, [Synset('association.n.08')], 'CASE 2:', 0, [], 'CASE 3:', 0, 'CASE 4:', 0, [], 'RULE:', 3]

['SENTENCE', 20678, ['data', 'information'], 'LOW', 'CASE 1:', 1, [Synset('datum.n.01')], 'CASE 2:', 0, [], 'CASE 3:', 0, 'CASE 4:', 0, [], 'RULE:', 3]

['SENTENCE', 23012, ['product', 'event'], 'LOW', 'CASE 1:', 1, [Synset('product.n.05')], 'CASE 2:', 0, [], 'CASE 3:', 0, 'CASE 4:', 0, [], 'RULE:', 3]

['SENTENCE', 26162, ['singer', 'musician'], 'LOW', 'CASE 1:', 1, [Synset('singer.n.01')], 'CASE 2:', 0, [], 'CASE 3:', 0, 'CASE 4:', 0, [], 'RULE:', 3]

['SENTENCE', 31197, ['country', 'region'], 'HIGH', 'CASE 1:', 1, [Synset('area.n.01')], 'CASE 2:', 0, [], 'CASE 3:', 0, 'CASE 4:', 0, [], 'RULE:', 3]

['SENTENCE', 33078, ['measure', 'activity'], 'LOW', 'CASE 1:', 1, [Synset('measurement.n.01')],  
 'CASE 2:', 0, [], 'CASE 3:', 0, 'CASE 4:', 0, [], 'RULE:', 3]  
 ['SENTENCE', 37505, ['time', 'period'], 'LOW', 'CASE 1:', 2, [Synset('time.n.02'),  
 Synset('time.n.03')], 'CASE 2:', 0, [], 'CASE 3:', 0, 'CASE 4:', 0, [], 'RULE:', 3]  
 ['SENTENCE', 44476, ['side', 'line'], 'HIGH', 'CASE 1:', 2, [Synset('side.n.07'),  
 Synset('side.n.08')], 'CASE 2:', 0, [], 'CASE 3:', 0, 'CASE 4:', 0, [], 'RULE:', 3]  
 ['SENTENCE', 51153, ['bond', 'security'], 'HIGH', 'CASE 1:', 1, [Synset('bond.n.02')], 'CASE 2:',  
 0, [], 'CASE 3:', 0, 'CASE 4:', 0, [], 'RULE:', 3]  
 ['SENTENCE', 59391, ['bridge', 'structure'], 'LOW', 'CASE 1:', 2, [Synset('bridge.n.01'),  
 Synset('bridge.n.04')], 'CASE 2:', 0, [], 'CASE 3:', 0, 'CASE 4:', 0, [], 'RULE:', 3]  
 ['SENTENCE', 59667, ['prosecution', 'action'], 'LOW', 'CASE 1:', 1, [Synset('prosecution.n.01')],  
 'CASE 2:', 0, [], 'CASE 3:', 0, 'CASE 4:', 0, [], 'RULE:', 3]  
 ['SENTENCE', 67339, ['spring', 'break'], 'LOW', 'CASE 1:', 1, [Synset('spring.v.05')], 'CASE 2:', 0,  
 [], 'CASE 3:', 0, 'CASE 4:', 0, [], 'RULE:', 3]  
 ['SENTENCE', 70790, ['loan', 'debt'], 'HIGH', 'CASE 1:', 1, [Synset('loan.n.01')], 'CASE 2:', 0, [],  
 'CASE 3:', 0, 'CASE 4:', 0, [], 'RULE:', 3]  
 ['SENTENCE', 797518, ['lace', 'fabric'], 'LOW', 'CASE 1:', 1, [Synset('lace.n.02')], 'CASE 2:', 0, [],  
 'CASE 3:', 0, 'CASE 4:', 0, [], 'RULE:', 3]  
 ['SENTENCE', 795746, ['attack', 'crime'], 'LOW', 'CASE 1:', 1, [Synset('attack.n.05')], 'CASE 2:',  
 0, [], 'CASE 3:', 0, 'CASE 4:', 0, [], 'RULE:', 3]  
 ['SENTENCE', 790134, ['Chief', 'leader'], 'LOW', 'CASE 1:', 2, [Synset('head.n.04'),  
 Synset('headman.n.02')], 'CASE 2:', 0, [], 'CASE 3:', 0, 'CASE 4:', 0, [], 'RULE:', 3]  
 ['SENTENCE', 785197, ['cap', 'control'], 'LOW', 'CASE 1:', 1, [Synset('ceiling.n.03')], 'CASE 2:',  
 0, [], 'CASE 3:', 0, 'CASE 4:', 0, [], 'RULE:', 3]

['SENTENCE', 118, ['paper', 'stuff'], 'LOW', 'CASE 1:', 1, [Synset('paper.n.01')], 'CASE 2:', 0, [],  
 'CASE 3:', 0, 'CASE 4:', 0, [], 'RULE:', 4]  
 ['SENTENCE', 14777, ['Medallion', 'honor'], 'MEDIUM', 'CASE 1:', 1, [Synset('decoration.n.02')],  
 'CASE 2:', 0, [], 'CASE 3:', 0, 'CASE 4:', 0, [], 'RULE:', 4]  
 ['SENTENCE', 23841, ['bombing', 'attack'], 'HIGH', 'CASE 1:', 2, [Synset('bombing.n.01'),  
 Synset('bombard.v.02')], 'CASE 2:', 0, [], 'CASE 3:', 0, 'CASE 4:', 0, [], 'RULE:', 4]  
 ['SENTENCE', 26723, ['salad', 'dish'], 'LOW', 'CASE 1:', 1, [Synset('salad.n.01')], 'CASE 2:', 0, [],  
 'CASE 3:', 0, 'CASE 4:', 0, [], 'RULE:', 4]  
 ['SENTENCE', 29803, ['venture', 'business'], 'LOW', 'CASE 1:', 1, [Synset('venture.n.03')], 'CASE  
 2:', 0, [], 'CASE 3:', 0, 'CASE 4:', 0, [], 'RULE:', 4]  
 ['SENTENCE', 43549, ['fraud', 'crime'], 'HIGH', 'CASE 1:', 1, [Synset('fraud.n.01')], 'CASE 2:', 0,  
 [], 'CASE 3:', 0, 'CASE 4:', 0, [], 'RULE:', 4]  
 ['SENTENCE', 47118, ['Google', 'search'], 'LOW', 'CASE 1:', 1, [Synset('google.v.01')], 'CASE  
 2:', 0, [], 'CASE 3:', 0, 'CASE 4:', 0, [], 'RULE:', 4]  
 ['SENTENCE', 54789, ['stint', 'job'], 'MEDIUM', 'CASE 1:', 1, [Synset('stint.n.03')], 'CASE 2:', 0, [],  
 'CASE 3:', 0, 'CASE 4:', 0, [], 'RULE:', 4]

['SENTENCE', 62489, ['job', 'responsibility'], 'LOW', 'CASE 1:', 1, [Synset('job.n.05')], 'CASE 2:', 0, [], 'CASE 3:', 0, 'CASE 4:', 0, [], 'RULE:', 4]  
 ['SENTENCE', 66877, ['Sea', 'water'], 'LOW', 'CASE 1:', 1, [Synset('sea.n.01')], 'CASE 2:', 0, [], 'CASE 3:', 0, 'CASE 4:', 0, [], 'RULE:', 4]  
 ['SENTENCE', 70221, ['gown', 'sleepwear'], 'LOW', 'CASE 1:', 1, [Synset('nightgown.n.01')], 'CASE 2:', 0, [], 'CASE 3:', 0, 'CASE 4:', 0, [], 'RULE:', 4]  
 ['SENTENCE', 75502, ['Citation', 'honor'], 'LOW', 'CASE 1:', 1, [Synset('citation.n.01')], 'CASE 2:', 0, [], 'CASE 3:', 0, 'CASE 4:', 0, [], 'RULE:', 4]  
 ['SENTENCE', 76471, ['store', 'outlet'], 'LOW', 'CASE 1:', 1, [Synset('shop.n.01')], 'CASE 2:', 0, [], 'CASE 3:', 0, 'CASE 4:', 0, [], 'RULE:', 4]  
 ['SENTENCE', 78409, ['Archdiocese', 'diocese'], 'LOW', 'CASE 1:', 1, [Synset('archdiocese.n.01')], 'CASE 2:', 0, [], 'CASE 3:', 0, 'CASE 4:', 0, [], 'RULE:', 4]  
 ['SENTENCE', 82975, ['burn', 'injury'], 'MEDIUM', 'CASE 1:', 1, [Synset('burn.n.03')], 'CASE 2:', 0, [], 'CASE 3:', 0, 'CASE 4:', 0, [], 'RULE:', 4]  
 ['SENTENCE', 88491, ['anger', 'emotion'], 'LOW', 'CASE 1:', 1, [Synset('anger.n.01')], 'CASE 2:', 0, [], 'CASE 3:', 0, 'CASE 4:', 0, [], 'RULE:', 4]  
 ['SENTENCE', 797537, ['floor', 'room'], 'LOW', 'CASE 1:', 1, [Synset('floor.n.10')], 'CASE 2:', 0, [], 'CASE 3:', 0, 'CASE 4:', 0, [], 'RULE:', 4]  
 ['SENTENCE', 793053, ['game', 'division'], 'LOW', 'CASE 1:', 1, [Synset('game.n.05')], 'CASE 2:', 0, [], 'CASE 3:', 0, 'CASE 4:', 0, [], 'RULE:', 4]  
 ['SENTENCE', 786726, ['researcher', 'scientist'], 'LOW', 'CASE 1:', 1, [Synset('research\_worker.n.01')], 'CASE 2:', 0, [], 'CASE 3:', 0, 'CASE 4:', 0, [], 'RULE:', 4]  
 ['SENTENCE', 750071, ['dealership', 'business'], 'LOW', 'CASE 1:', 1, [Synset('franchise.n.02')], 'CASE 2:', 0, [], 'CASE 3:', 0, 'CASE 4:', 0, [], 'RULE:', 4]  
  
 ['SENTENCE', 16200, ['change', 'change'], 'LOW', 'CASE 1:', 1, [Synset('deepen.v.04')], 'CASE 2:', 1, [Synset('deepen.v.04')], 'CASE 3:', 0, 'CASE 4:', 0, [], 'RULE:', 5]  
 ['SENTENCE', 154888, ['cotton', 'material'], 'LOW', 'CASE 1:', 1, [Synset('cotton.n.02')], 'CASE 2:', 0, [], 'CASE 3:', 0, 'CASE 4:', 0, [], 'RULE:', 5]  
 ['SENTENCE', 620934, ['pudding', 'Desserts'], 'LOW', 'CASE 1:', 2, [Synset('pudding.n.02'), Synset('pudding.n.03')], 'CASE 2:', 0, [], 'CASE 3:', 0, 'CASE 4:', 0, [], 'RULE:', 5]  
 ['SENTENCE', 658853, ['Avenue', 'street'], 'LOW', 'CASE 1:', 1, [Synset('avenue.n.02')], 'CASE 2:', 0, [], 'CASE 3:', 0, 'CASE 4:', 0, [], 'RULE:', 5]  
 ['SENTENCE', 701930, ['way', 'condition'], 'LOW', 'CASE 1:', 1, [Synset('way.n.04')], 'CASE 2:', 0, [], 'CASE 3:', 0, 'CASE 4:', 0, [], 'RULE:', 5]  
 ['SENTENCE', 716240, ['railroad', 'force'], 'LOW', 'CASE 1:', 1, [Synset('dragoon.v.01')], 'CASE 2:', 0, [], 'CASE 3:', 0, 'CASE 4:', 0, [], 'RULE:', 5]  
 ['SENTENCE', 770884, ['hero', 'character'], 'LOW', 'CASE 1:', 1, [Synset('hero.n.02')], 'CASE 2:', 0, [], 'CASE 3:', 0, 'CASE 4:', 0, [], 'RULE:', 5]

CASE 2

['SENTENCE', 310366, ['name', 'brand'], 'LOW', 'CASE 1:', 0, [], 'CASE 2:', 1, [Synset('trade\_name.n.01')], 'CASE 3:', 1, 'CASE 4:', 0, [], 'RULE:', 1]

['SENTENCE', 211582, ['Press', 'paper'], 'LOW', 'CASE 1:', 0, [], 'CASE 2:', 1, [Synset('newspaper.n.01')], 'CASE 3:', 1, 'CASE 4:', 0, [], 'RULE:', 2]

['SENTENCE', 352357, ['air', 'issue'], 'LOW', 'CASE 1:', 0, [], 'CASE 2:', 1, [Synset('publish.v.02')], 'CASE 3:', 1, 'CASE 4:', 0, [], 'RULE:', 2]

['SENTENCE', 2526, ['debt', 'obligation'], 'LOW', 'CASE 1:', 1, [Synset('debt.n.03')], 'CASE 2:', 0, [], 'CASE 3:', 0, 'CASE 4:', 0, [], 'RULE:', 3]

['SENTENCE', 15093, ['lyric', 'writing'], 'MEDIUM', 'CASE 1:', 1, [Synset('lyric.v.01')], 'CASE 2:', 0, [], 'CASE 3:', 0, 'CASE 4:', 0, [], 'RULE:', 3]

['SENTENCE', 20358, ['association', 'group'], 'MEDIUM', 'CASE 1:', 1, [Synset('association.n.08')], 'CASE 2:', 0, [], 'CASE 3:', 0, 'CASE 4:', 0, [], 'RULE:', 3]

['SENTENCE', 20678, ['data', 'information'], 'LOW', 'CASE 1:', 1, [Synset('datum.n.01')], 'CASE 2:', 0, [], 'CASE 3:', 0, 'CASE 4:', 0, [], 'RULE:', 3]

['SENTENCE', 23012, ['product', 'event'], 'LOW', 'CASE 1:', 1, [Synset('product.n.05')], 'CASE 2:', 0, [], 'CASE 3:', 0, 'CASE 4:', 0, [], 'RULE:', 3]

['SENTENCE', 26162, ['singer', 'musician'], 'LOW', 'CASE 1:', 1, [Synset('singer.n.01')], 'CASE 2:', 0, [], 'CASE 3:', 0, 'CASE 4:', 0, [], 'RULE:', 3]

['SENTENCE', 31197, ['country', 'region'], 'HIGH', 'CASE 1:', 1, [Synset('area.n.01')], 'CASE 2:', 0, [], 'CASE 3:', 0, 'CASE 4:', 0, [], 'RULE:', 3]

['SENTENCE', 33078, ['measure', 'activity'], 'LOW', 'CASE 1:', 1, [Synset('measurement.n.01')], 'CASE 2:', 0, [], 'CASE 3:', 0, 'CASE 4:', 0, [], 'RULE:', 3]

['SENTENCE', 37505, ['time', 'period'], 'LOW', 'CASE 1:', 2, [Synset('time.n.02'), Synset('time.n.03')], 'CASE 2:', 0, [], 'CASE 3:', 0, 'CASE 4:', 0, [], 'RULE:', 3]

['SENTENCE', 44476, ['side', 'line'], 'HIGH', 'CASE 1:', 2, [Synset('side.n.07'), Synset('side.n.08')], 'CASE 2:', 0, [], 'CASE 3:', 0, 'CASE 4:', 0, [], 'RULE:', 3]

['SENTENCE', 51153, ['bond', 'security'], 'HIGH', 'CASE 1:', 1, [Synset('bond.n.02')], 'CASE 2:', 0, [], 'CASE 3:', 0, 'CASE 4:', 0, [], 'RULE:', 3]

['SENTENCE', 59391, ['bridge', 'structure'], 'LOW', 'CASE 1:', 2, [Synset('bridge.n.01'), Synset('bridge.n.04')], 'CASE 2:', 0, [], 'CASE 3:', 0, 'CASE 4:', 0, [], 'RULE:', 3]

['SENTENCE', 59667, ['prosecution', 'action'], 'LOW', 'CASE 1:', 1, [Synset('prosecution.n.01')], 'CASE 2:', 0, [], 'CASE 3:', 0, 'CASE 4:', 0, [], 'RULE:', 3]

['SENTENCE', 67339, ['spring', 'break'], 'LOW', 'CASE 1:', 1, [Synset('spring.v.05')], 'CASE 2:', 0, [], 'CASE 3:', 0, 'CASE 4:', 0, [], 'RULE:', 3]

['SENTENCE', 70790, ['loan', 'debt'], 'HIGH', 'CASE 1:', 1, [Synset('loan.n.01')], 'CASE 2:', 0, [], 'CASE 3:', 0, 'CASE 4:', 0, [], 'RULE:', 3]

['SENTENCE', 76435, ['shoplifter', 'thief'], 'LOW', 'CASE 1:', 1, [Synset('booster.n.03')], 'CASE 2:', 0, [], 'CASE 3:', 0, 'CASE 4:', 0, [], 'RULE:', 3]

['SENTENCE', 81233, ['cap', 'covering'], 'LOW', 'CASE 1:', 1, [Synset('hood.n.02')], 'CASE 2:', 0, [], 'CASE 3:', 0, 'CASE 4:', 0, [], 'RULE:', 3]

['SENTENCE', 84035, ['woman', 'class'], 'LOW', 'CASE 1:', 1, [Synset('womanhood.n.02')],  
 'CASE 2:', 0, [], 'CASE 3:', 0, 'CASE 4:', 0, [], 'RULE:', 3]  
 ['SENTENCE', 90665, ['violation', 'crime'], 'LOW', 'CASE 1:', 1, [Synset('misdemeanor.n.01')],  
 'CASE 2:', 0, [], 'CASE 3:', 0, 'CASE 4:', 0, [], 'RULE:', 3]  
 ['SENTENCE', 99658, ['finalist', 'competition'], 'LOW', 'CASE 1:', 1, [Synset('finalist.n.01')],  
 'CASE 2:', 0, [], 'CASE 3:', 0, 'CASE 4:', 0, [], 'RULE:', 3]

['SENTENCE', 12245, ['chemical', 'fertilizer'], 'MEDIUM', 'CASE 1:', 0, [], 'CASE 2:', 1,  
 [Synset('fertilizer.n.01')], 'CASE 3:', 1, 'CASE 4:', 0, [], 'RULE:', 4]  
 ['SENTENCE', 60229, ['trip', 'travel'], 'LOW', 'CASE 1:', 0, [], 'CASE 2:', 1, [Synset('travel.v.02')],  
 'CASE 3:', 1, 'CASE 4:', 0, [], 'RULE:', 4]  
 ['SENTENCE', 72539, ['affiliate', 'company'], 'LOW', 'CASE 1:', 0, [], 'CASE 2:', 1,  
 [Synset('company.v.01')], 'CASE 3:', 1, 'CASE 4:', 0, [], 'RULE:', 4]  
 ['SENTENCE', 98258, ['effort', 'reform'], 'LOW', 'CASE 1:', 0, [], 'CASE 2:', 1,  
 [Synset('reform.n.02')], 'CASE 3:', 1, 'CASE 4:', 0, [], 'RULE:', 4]  
 ['SENTENCE', 123965, ['status', 'way'], 'LOW', 'CASE 1:', 0, [], 'CASE 2:', 1, [Synset('way.n.04')],  
 'CASE 3:', 1, 'CASE 4:', 0, [], 'RULE:', 4]  
 ['SENTENCE', 131102, ['property', 'estate'], 'LOW', 'CASE 1:', 0, [], 'CASE 2:', 1,  
 [Synset('estate.n.01')], 'CASE 3:', 1, 'CASE 4:', 0, [], 'RULE:', 4]  
 ['SENTENCE', 139468, ['sketch', 'design'], 'LOW', 'CASE 1:', 0, [], 'CASE 2:', 1,  
 [Synset('design.n.06')], 'CASE 3:', 1, 'CASE 4:', 0, [], 'RULE:', 4]  
 ['SENTENCE', 148712, ['television', 'cable'], 'LOW', 'CASE 1:', 0, [], 'CASE 2:', 2,  
 [Synset('cable\_television.n.01'), Synset('cable.n.06')], 'CASE 3:', 1, 'CASE 4:', 0, [], 'RULE:', 4]  
 ['SENTENCE', 151181, ['machine', 'lever'], 'MEDIUM', 'CASE 1:', 0, [], 'CASE 2:', 1,  
 [Synset('lever.n.02')], 'CASE 3:', 1, 'CASE 4:', 0, [], 'RULE:', 4]  
 ['SENTENCE', 179593, ['institution', 'charity'], 'LOW', 'CASE 1:', 0, [], 'CASE 2:', 1,  
 [Synset('charity.n.05')], 'CASE 3:', 1, 'CASE 4:', 0, [], 'RULE:', 4]  
 ['SENTENCE', 193252, ['line', 'rope'], 'LOW', 'CASE 1:', 0, [], 'CASE 2:', 1, [Synset('rope.n.01')],  
 'CASE 3:', 1, 'CASE 4:', 0, [], 'RULE:', 4]  
 ['SENTENCE', 226915, ['tip', 'head'], 'LOW', 'CASE 1:', 0, [], 'CASE 2:', 1, [Synset('head.n.10')],  
 'CASE 3:', 1, 'CASE 4:', 0, [], 'RULE:', 4]  
 ['SENTENCE', 227301, ['gathering', 'community'], 'LOW', 'CASE 1:', 0, [], 'CASE 2:', 1,  
 [Synset('community.n.01')], 'CASE 3:', 1, 'CASE 4:', 0, [], 'RULE:', 4]  
 ['SENTENCE', 239662, ['trade', 'deal'], 'LOW', 'CASE 1:', 0, [], 'CASE 2:', 1, [Synset('deal.v.13')],  
 'CASE 3:', 1, 'CASE 4:', 0, [], 'RULE:', 4]  
 ['SENTENCE', 277045, ['show', 'program'], 'LOW', 'CASE 1:', 0, [], 'CASE 2:', 1,  
 [Synset('broadcast.n.02')], 'CASE 3:', 1, 'CASE 4:', 0, [], 'RULE:', 4]  
 ['SENTENCE', 290057, ['attack', 'bombing'], 'MEDIUM', 'CASE 1:', 0, [], 'CASE 2:', 2,  
 [Synset('bombing.n.01'), Synset('bombard.v.02')], 'CASE 3:', 1, 'CASE 4:', 0, [], 'RULE:', 4]  
 ['SENTENCE', 323939, ['placement', 'intervention'], 'LOW', 'CASE 1:', 0, [], 'CASE 2:', 1,  
 [Synset('interposition.n.02')], 'CASE 3:', 1, 'CASE 4:', 0, [], 'RULE:', 4]  
 ['SENTENCE', 347538, ['party', 'wedding'], 'LOW', 'CASE 1:', 0, [], 'CASE 2:', 1,  
 [Synset('wedding.n.03')], 'CASE 3:', 1, 'CASE 4:', 0, [], 'RULE:', 4]

['SENTENCE', 375468, ['worker', 'employee'], 'MEDIUM', 'CASE 1:', 0, [], 'CASE 2:', 1, [Synset('employee.n.01')], 'CASE 3:', 1, 'CASE 4:', 0, [], 'RULE:', 4]  
['SENTENCE', 381330, ['payment', 'bonus'], 'LOW', 'CASE 1:', 0, [], 'CASE 2:', 1, [Synset('bonus.n.02')], 'CASE 3:', 1, 'CASE 4:', 0, [], 'RULE:', 4]  
['SENTENCE', 389423, ['being', 'answer'], 'LOW', 'CASE 1:', 0, [], 'CASE 2:', 1, [Synset('answer.v.06')], 'CASE 3:', 1, 'CASE 4:', 0, [], 'RULE:', 4]

['SENTENCE', 192757, ['player', 'star'], 'LOW', 'CASE 1:', 0, [], 'CASE 2:', 1, [Synset('star.n.04')], 'CASE 3:', 1, 'CASE 4:', 0, [], 'RULE:', 5]  
['SENTENCE', 228927, ['program', 'budget'], 'LOW', 'CASE 1:', 0, [], 'CASE 2:', 1, [Synset('budget.n.02')], 'CASE 3:', 1, 'CASE 4:', 0, [], 'RULE:', 5]  
['SENTENCE', 545416, ['maker', 'manufacturer'], 'LOW', 'CASE 1:', 0, [], 'CASE 2:', 1, [Synset('manufacturer.n.02')], 'CASE 3:', 1, 'CASE 4:', 0, [], 'RULE:', 5]  
['SENTENCE', 549001, ['number', 'record'], 'LOW', 'CASE 1:', 0, [], 'CASE 2:', 1, [Synset('record.n.03')], 'CASE 3:', 1, 'CASE 4:', 0, [], 'RULE:', 5]  
['SENTENCE', 631155, ['juvenile', 'minor'], 'LOW', 'CASE 1:', 0, [], 'CASE 2:', 1, [Synset('child.n.01')], 'CASE 3:', 1, 'CASE 4:', 0, [], 'RULE:', 5]  
['SENTENCE', 738814, ['advocate', 'Democrats'], 'LOW', 'CASE 1:', 0, [], 'CASE 2:', 1, [Synset('democrat.n.02')], 'CASE 3:', 1, 'CASE 4:', 0, [], 'RULE:', 5]

### CASE 3

['SENTENCE', 4259, ['stock', 'asset'], 'HIGH', 'CASE 1:', 0, [], 'CASE 2:', 0, [], 'CASE 3:', 1, 'CASE 4:', 0, [], 'RULE:', 1]  
CORRECT

['SENTENCE', 471844, ['Duchamp', 'artist'], 'HIGH', 'CASE 1:', 0, [], 'CASE 2:', 0, [], 'CASE 3:', 1, 'CASE 4:', 0, [], 'RULE:', 1]  
CORRECT

['SENTENCE', 552546, ['sharing', 'file'], 'HIGH', 'CASE 1:', 0, [], 'CASE 2:', 0, [], 'CASE 3:', 1, 'CASE 4:', 0, [], 'RULE:', 1]  
CORRECT

['SENTENCE', 25618, ['Island', 'area'], 'MEDIUM', 'CASE 1:', 0, [], 'CASE 2:', 0, [], 'CASE 3:', 1, 'CASE 4:', 0, [], 'RULE:', 1]  
CORRECT

['SENTENCE', 145601, ['family', 'institution'], 'MEDIUM', 'CASE 1:', 0, [], 'CASE 2:', 0, [], 'CASE 3:', 1, 'CASE 4:', 0, [], 'RULE:', 1]  
INCORRECT

['SENTENCE', 249081, ['show', 'television'], 'MEDIUM', 'CASE 1:', 0, [], 'CASE 2:', 0, [], 'CASE 3:', 1, 'CASE 4:', 0, [], 'RULE:', 1]  
CORRECT

['SENTENCE', 249081, ['Law', 'television'], 'MEDIUM', 'CASE 1:', 0, [], 'CASE 2:', 0, [], 'CASE 3:', 1, 'CASE 4:', 0, [], 'RULE:', 1]  
INCORRECT

['SENTENCE', 549595, ['Old', 'restaurant'], 'LOW', 'CASE 1:', 0, [], 'CASE 2:', 0, [], 'CASE 3:', 1, 'CASE 4:', 0, [], 'RULE:', 1]  
UNCERTAIN

['SENTENCE', 549595, ['Homestead', 'restaurant'], 'LOW', 'CASE 1:', 0, [], 'CASE 2:', 0, [], 'CASE 3:', 1, 'CASE 4:', 0, [], 'RULE:', 1]  
INCORRECT

['SENTENCE', 548607, ['Spirit', 'group'], 'LOW', 'CASE 1:', 0, [], 'CASE 2:', 0, [], 'CASE 3:', 1, 'CASE 4:', 0, [], 'RULE:', 1]  
CORRECT

['SENTENCE', 3121, ['drug', 'issue'], 'MEDIUM', 'CASE 1:', 0, [], 'CASE 2:', 0, [], 'CASE 3:', 1, 'CASE 4:', 0, [], 'RULE:', 2]  
UNCERTAIN

['SENTENCE', 3121, ['benefit', 'issue'], 'MEDIUM', 'CASE 1:', 0, [], 'CASE 2:', 0, [], 'CASE 3:', 1, 'CASE 4:', 0, [], 'RULE:', 2]  
CORRECT

['SENTENCE', 32854, ['service', 'thing'], 'MEDIUM', 'CASE 1:', 0, [], 'CASE 2:', 0, [], 'CASE 3:', 1, 'CASE 4:', 0, [], 'RULE:', 2]  
UNCERTAIN

['SENTENCE', 99755, ['family', 'estate'], 'MEDIUM', 'CASE 1:', 0, [], 'CASE 2:', 0, [], 'CASE 3:', 1, 'CASE 4:', 0, [], 'RULE:', 2]  
CORRECT

['SENTENCE', 178482, ['game', 'thing'], 'MEDIUM', 'CASE 1:', 0, [], 'CASE 2:', 0, [], 'CASE 3:', 1, 'CASE 4:', 0, [], 'RULE:', 2]  
INCORRECT

['SENTENCE', 206370, ['Sterling', 'contemporary'], 'HIGH', 'CASE 1:', 0, [], 'CASE 2:', 0, [], 'CASE 3:', 1, 'CASE 4:', 0, [], 'RULE:', 2]  
INCORRECT

['SENTENCE', 206370, ['Ruby', 'contemporary'], 'HIGH', 'CASE 1:', 0, [], 'CASE 2:', 0, [], 'CASE 3:', 1, 'CASE 4:', 0, [], 'RULE:', 2]  
INCORRECT

['SENTENCE', 102, ['I', 'tune'], 'LOW', 'CASE 1:', 0, [], 'CASE 2:', 0, [], 'CASE 3:', 1, 'CASE 4:', 0, [], 'RULE:', 2]  
INCORRECT

['SENTENCE', 102, ['Sniff', 'tune'], 'LOW', 'CASE 1:', 0, [], 'CASE 2:', 0, [], 'CASE 3:', 1, 'CASE 4:', 0, [], 'RULE:', 2]  
INCORRECT

['SENTENCE', 1757, ['pool', 'creation'], 'LOW', 'CASE 1:', 0, [], 'CASE 2:', 0, [], 'CASE 3:', 1, 'CASE 4:', 0, [], 'RULE:', 2]  
UNCERTAIN

['SENTENCE', 1048, ['Iraq', 'nation'], 'HIGH', 'CASE 1:', 0, [], 'CASE 2:', 0, [], 'CASE 3:', 1, 'CASE 4:', 0, [], 'RULE:', 3]  
CORRECT

['SENTENCE', 2521, ['bank', 'institution'], 'HIGH', 'CASE 1:', 0, [], 'CASE 2:', 0, [], 'CASE 3:', 1, 'CASE 4:', 0, [], 'RULE:', 3]  
CORRECT

['SENTENCE', 2947, ['family', 'member'], 'HIGH', 'CASE 1:', 0, [], 'CASE 2:', 0, [], 'CASE 3:', 1, 'CASE 4:', 0, [], 'RULE:', 3]  
CORRECT

['SENTENCE', 3193, ['tax', 'issue'], 'HIGH', 'CASE 1:', 0, [], 'CASE 2:', 0, [], 'CASE 3:', 1, 'CASE 4:', 0, [], 'RULE:', 3]  
UNCERTAIN

['SENTENCE', 316, ['Mr.', 'owner'], 'MEDIUM', 'CASE 1:', 0, [], 'CASE 2:', 0, [], 'CASE 3:', 1, 'CASE 4:', 0, [], 'RULE:', 3]  
INCORRECT

['SENTENCE', 1696, ['station', 'area'], 'MEDIUM', 'CASE 1:', 0, [], 'CASE 2:', 0, [], 'CASE 3:', 1, 'CASE 4:', 0, [], 'RULE:', 3]  
UNCERTAIN

['SENTENCE', 2283, ['country', 'Iran'], 'MEDIUM', 'CASE 1:', 0, [], 'CASE 2:', 0, [], 'CASE 3:', 1, 'CASE 4:', 0, [], 'RULE:', 3]



CORRECT

['SENTENCE', 2346, ['reality', 'scene'], 'MEDIUM', 'CASE 1:', 0, [], 'CASE 2:', 0, [], 'CASE 3:', 1, 'CASE 4:', 0, [], 'RULE:', 3]

INCORRECT

['SENTENCE', 212, ['reform', 'country'], 'LOW', 'CASE 1:', 0, [], 'CASE 2:', 0, [], 'CASE 3:', 1, 'CASE 4:', 0, [], 'RULE:', 3]

INCORRECT

['SENTENCE', 302, ['time', 'mile'], 'LOW', 'CASE 1:', 0, [], 'CASE 2:', 0, [], 'CASE 3:', 1, 'CASE 4:', 0, [], 'RULE:', 3]

INCORRECT

['SENTENCE', 1720, ['station', 'dozen'], 'HIGH', 'CASE 1:', 0, [], 'CASE 2:', 0, [], 'CASE 3:', 1, 'CASE 4:', 0, [], 'RULE:', 4]

INCORRECT

['SENTENCE', 3060, ['Washington', 'city'], 'HIGH', 'CASE 1:', 0, [], 'CASE 2:', 0, [], 'CASE 3:', 1, 'CASE 4:', 0, [], 'RULE:', 4]

INCORRECT

['SENTENCE', 4171, ['executive', 'company'], 'HIGH', 'CASE 1:', 0, [], 'CASE 2:', 0, [], 'CASE 3:', 1, 'CASE 4:', 0, [], 'RULE:', 4]

CORRECT

['SENTENCE', 4791, ['article', 'material'], 'HIGH', 'CASE 1:', 0, [], 'CASE 2:', 0, [], 'CASE 3:', 1, 'CASE 4:', 0, [], 'RULE:', 4]

CORRECT

['SENTENCE', 5846, ['Bureau', 'agency'], 'MEDIUM', 'CASE 1:', 0, [], 'CASE 2:', 0, [], 'CASE 3:', 1, 'CASE 4:', 0, [], 'RULE:', 4]

CORRECT

['SENTENCE', 6671, ['performance', 'event'], 'MEDIUM', 'CASE 1:', 0, [], 'CASE 2:', 0, [], 'CASE 3:', 1, 'CASE 4:', 0, [], 'RULE:', 4]

CORRECT

['SENTENCE', 6953, ['example', 'work'], 'MEDIUM', 'CASE 1:', 0, [], 'CASE 2:', 0, [], 'CASE 3:', 1, 'CASE 4:', 0, [], 'RULE:', 4]

INCORRECT

['SENTENCE', 106, ['Theresa', 'archive'], 'LOW', 'CASE 1:', 0, [], 'CASE 2:', 0, [], 'CASE 3:', 1, 'CASE 4:', 0, [], 'RULE:', 4]  
INCORRECT

['SENTENCE', 106, ['Stern', 'archive'], 'LOW', 'CASE 1:', 0, [], 'CASE 2:', 0, [], 'CASE 3:', 1, 'CASE 4:', 0, [], 'RULE:', 4]  
INCORRECT

['SENTENCE', 1155, ['Hezbollah', 'organization'], 'LOW', 'CASE 1:', 0, [], 'CASE 2:', 0, [], 'CASE 3:', 1, 'CASE 4:', 0, [], 'RULE:', 4]  
INCORRECT

['SENTENCE', 11707, ['woman', 'people'], 'HIGH', 'CASE 1:', 0, [], 'CASE 2:', 0, [], 'CASE 3:', 1, 'CASE 4:', 0, [], 'RULE:', 5]  
CORRECT

['SENTENCE', 77967, ['Brazil', 'country'], 'HIGH', 'CASE 1:', 0, [], 'CASE 2:', 0, [], 'CASE 3:', 1, 'CASE 4:', 0, [], 'RULE:', 5]  
CORRECT

['SENTENCE', 412097, ['health', 'issue'], 'HIGH', 'CASE 1:', 0, [], 'CASE 2:', 0, [], 'CASE 3:', 1, 'CASE 4:', 0, [], 'RULE:', 5]  
UNCERTAIN

['SENTENCE', 13639, ['energy', 'material'], 'MEDIUM', 'CASE 1:', 0, [], 'CASE 2:', 0, [], 'CASE 3:', 1, 'CASE 4:', 0, [], 'RULE:', 5]  
UNCERTAIN

['SENTENCE', 34745, ['men', 'people'], 'MEDIUM', 'CASE 1:', 0, [], 'CASE 2:', 0, [], 'CASE 3:', 1, 'CASE 4:', 0, [], 'RULE:', 5]  
CORRECT

['SENTENCE', 54946, ['woman', 'voter'], 'MEDIUM', 'CASE 1:', 0, [], 'CASE 2:', 0, [], 'CASE 3:', 1, 'CASE 4:', 0, [], 'RULE:', 5]  
UNCERTAIN

['SENTENCE', 71134, ['girl', 'teenager'], 'MEDIUM', 'CASE 1:', 0, [], 'CASE 2:', 0, [], 'CASE 3:', 1, 'CASE 4:', 0, [], 'RULE:', 5]  
CORRECT

['SENTENCE', 1774, ['people', 'talent'], 'LOW', 'CASE 1:', 0, [], 'CASE 2:', 0, [], 'CASE 3:', 1, 'CASE 4:', 0, [], 'RULE:', 5]  
INCORRECT

['SENTENCE', 23416, ['Jews', 'group'], 'LOW', 'CASE 1:', 0, [], 'CASE 2:', 0, [], 'CASE 3:', 1, 'CASE 4:', 0, [], 'RULE:', 5]  
CORRECT

['SENTENCE', 25759, ['flavor', 'fruit'], 'LOW', 'CASE 1:', 0, [], 'CASE 2:', 0, [], 'CASE 3:', 1, 'CASE 4:', 0, [], 'RULE:', 5]  
CORRECT

CASE 4

['SENTENCE', 765718, [""multicultural", 'expression'], 'LOW', 'CASE 1:', 0, [], 'CASE 2:', 0, [], 'CASE 3:', 1, 'CASE 4:', 1, [""multicultural"], 'RULE:', 1]  
INCORRECT

['SENTENCE', 755901, ['Khodorkovsky', 'prosecution'], 'LOW', 'CASE 1:', 0, [], 'CASE 2:', 0, [], 'CASE 3:', 1, 'CASE 4:', 1, ['Khodorkovsky'], 'RULE:', 1]  
INCORRECT

['SENTENCE', 724833, ['Arkadius', 'Britain'], 'LOW', 'CASE 1:', 0, [], 'CASE 2:', 0, [], 'CASE 3:', 1, 'CASE 4:', 1, ['Arkadius'], 'RULE:', 1]  
INCORRECT

['SENTENCE', 684178, ['ribeye', 'flourish'], 'LOW', 'CASE 1:', 0, [], 'CASE 2:', 0, [], 'CASE 3:', 1, 'CASE 4:', 1, ['ribeye'], 'RULE:', 1]  
INCORRECT

['SENTENCE', 573768, ['Roadshow', 'television'], 'LOW', 'CASE 1:', 0, [], 'CASE 2:', 0, [], 'CASE 3:', 1, 'CASE 4:', 1, ['Roadshow'], 'RULE:', 1]  
CORRECT

['SENTENCE', 573408, ['Bolognese', 'food'], 'LOW', 'CASE 1:', 0, [], 'CASE 2:', 0, [], 'CASE 3:', 1, 'CASE 4:', 1, ['Bolognese'], 'RULE:', 1]  
CORRECT

['SENTENCE', 500698, [""64A7BFCFC3104A0B98841C3545", 'F.'], 'LOW', 'CASE 1:', 0, [], 'CASE 2:', 0, [], 'CASE 3:', 1, 'CASE 4:', 1, [""64A7BFCFC3104A0B98841C3545", 'F.'], 'RULE:', 1]  
INCORRECT

['SENTENCE', 475296, ['vehicle', 'off-road'], 'LOW', 'CASE 1:', 0, [], 'CASE 2:', 0, [], 'CASE 3:', 1, 'CASE 4:', 1, ['off-road'], 'RULE:', 1]  
CORRECT

['SENTENCE', 376377, ['Mobileistic', 'master'], 'LOW', 'CASE 1:', 0, [], 'CASE 2:', 0, [], 'CASE 3:', 1, 'CASE 4:', 1, ['Mobileistic'], 'RULE:', 1]  
INCORRECT

['SENTENCE', 314949, ['housewares', 'product'], 'LOW', 'CASE 1:', 0, [], 'CASE 2:', 0, [], 'CASE 3:', 1, 'CASE 4:', 1, ['housewares'], 'RULE:', 1]  
CORRECT

['SENTENCE', 102, ['Wanna', 'tune'], 'LOW', 'CASE 1:', 0, [], 'CASE 2:', 0, [], 'CASE 3:', 1, 'CASE 4:', 1, ['Wanna'], 'RULE:', 2]  
INCORRECT

['SENTENCE', 717820, ['Intracoastal', 'thing'], 'LOW', 'CASE 1:', 0, [], 'CASE 2:', 0, [], 'CASE 3:', 1, 'CASE 4:', 1, ['Intracoastal'], 'RULE:', 2]  
INCORRECT

['SENTENCE', 662052, ['proliferators', 'sale'], 'LOW', 'CASE 1:', 0, [], 'CASE 2:', 0, [], 'CASE 3:', 1, 'CASE 4:', 1, ['proliferators'], 'RULE:', 2]  
UNCERTAIN

['SENTENCE', 621139, ['bobotie', 'dish'], 'LOW', 'CASE 1:', 0, [], 'CASE 2:', 0, [], 'CASE 3:', 1, 'CASE 4:', 1, ['bobotie'], 'RULE:', 2]  
INCORRECT

['SENTENCE', 601453, ['recording', 'consumery'], 'LOW', 'CASE 1:', 0, [], 'CASE 2:', 0, [], 'CASE 3:', 1, 'CASE 4:', 1, ['consumery'], 'RULE:', 2]  
INCORRECT

['SENTENCE', 504843, ['Street', 'standouts'], 'LOW', 'CASE 1:', 0, [], 'CASE 2:', 0, [], 'CASE 3:', 1, 'CASE 4:', 1, ['standouts'], 'RULE:', 2]  
CORRECT

['SENTENCE', 501743, ['reweaves', 'feature'], 'LOW', 'CASE 1:', 0, [], 'CASE 2:', 0, [], 'CASE 3:', 1, 'CASE 4:', 1, ['reweaves'], 'RULE:', 2]  
INCORRECT

['SENTENCE', 211582, ['Newsday', 'paper'], 'LOW', 'CASE 1:', 0, [], 'CASE 2:', 0, [], 'CASE 3:', 1, 'CASE 4:', 1, ['Newsday'], 'RULE:', 2]  
CORRECT

['SENTENCE', 211582, ['Observateur', 'paper'], 'LOW', 'CASE 1:', 0, [], 'CASE 2:', 0, [], 'CASE 3:', 1, 'CASE 4:', 1, ['Observateur'], 'RULE:', 2]

CORRECT

['SENTENCE', 217199, ['self-deluders', 'posturing'], 'LOW', 'CASE 1:', 0, [], 'CASE 2:', 0, [], 'CASE 3:', 1, 'CASE 4:', 1, ['self-deluders'], 'RULE:', 2]

UNCERTAIN

['SENTENCE', 255510, ['plant-life', 'medium'], 'MEDIUM', 'CASE 1:', 0, [], 'CASE 2:', 0, [], 'CASE 3:', 1, 'CASE 4:', 1, ['plant-life'], 'RULE:', 3]

INCORRECT

['SENTENCE', 316, ['small-business', 'owner'], 'LOW', 'CASE 1:', 0, [], 'CASE 2:', 0, [], 'CASE 3:', 1, 'CASE 4:', 1, ['small-business'], 'RULE:', 3]

CORRECT

['SENTENCE', 8810, ['high-value', 'target'], 'LOW', 'CASE 1:', 0, [], 'CASE 2:', 0, [], 'CASE 3:', 1, 'CASE 4:', 1, ['high-value'], 'RULE:', 3]

CORRECT

['SENTENCE', 16130, ['Evangelicals', 'conservative'], 'LOW', 'CASE 1:', 0, [], 'CASE 2:', 0, [], 'CASE 3:', 1, 'CASE 4:', 1, ['Evangelicals'], 'RULE:', 3]

UNCERTAIN

['SENTENCE', 110366, ['defenseman', 'history'], 'HIGH', 'CASE 1:', 0, [], 'CASE 2:', 0, [], 'CASE 3:', 1, 'CASE 4:', 1, ['defenseman'], 'RULE:', 3]

UNCERTAIN

['SENTENCE', 799208, ['low-energy', 'task'], 'LOW', 'CASE 1:', 0, [], 'CASE 2:', 0, [], 'CASE 3:', 1, 'CASE 4:', 1, ['low-energy'], 'RULE:', 3]

CORRECT

['SENTENCE', 798506, ['Microsoft', 'product'], 'MEDIUM', 'CASE 1:', 0, [], 'CASE 2:', 0, [], 'CASE 3:', 1, 'CASE 4:', 1, ['Microsoft'], 'RULE:', 3]

UNCERTAIN

['SENTENCE', 791963, ['cassoulet', 'meat'], 'LOW', 'CASE 1:', 0, [], 'CASE 2:', 0, [], 'CASE 3:', 1, 'CASE 4:', 1, ['cassoulet'], 'RULE:', 3]

CORRECT

['SENTENCE', 781766, ['broker', 'cyber'], 'LOW', 'CASE 1:', 0, [], 'CASE 2:', 0, [], 'CASE 3:', 1, 'CASE 4:', 1, ['cyber'], 'RULE:', 3]

UNCERTAIN

['SENTENCE', 778909, ['something', 'slickness'], 'LOW', 'CASE 1:', 0, [], 'CASE 2:', 0, [], 'CASE 3:', 1, 'CASE 4:', 1, ['something'], 'RULE:', 3]

['SENTENCE', 7224, ['feature', 'imax'], 'HIGH', 'CASE 1:', 0, [], 'CASE 2:', 0, [], 'CASE 3:', 1, 'CASE 4:', 1, ['imax'], 'RULE:', 4]  
CORRECT

['SENTENCE', 60526, ['Citigroup', 'bank'], 'HIGH', 'CASE 1:', 0, [], 'CASE 2:', 0, [], 'CASE 3:', 1, 'CASE 4:', 1, ['Citigroup'], 'RULE:', 4]  
CORRECT

['SENTENCE', 108009, ['program', 'NBC'], 'MEDIUM', 'CASE 1:', 0, [], 'CASE 2:', 0, [], 'CASE 3:', 1, 'CASE 4:', 1, ['NBC'], 'RULE:', 4]  
CORRECT

['SENTENCE', 119549, [""'Funny", 'film'], 'MEDIUM', 'CASE 1:', 0, [], 'CASE 2:', 0, [], 'CASE 3:', 1, 'CASE 4:', 1, [""'Funny"], 'RULE:', 4]  
UNCERTAIN

['SENTENCE', 799446, ['schoolchildren', 'people'], 'LOW', 'CASE 1:', 0, [], 'CASE 2:', 0, [], 'CASE 3:', 1, 'CASE 4:', 1, ['schoolchildren'], 'RULE:', 4]  
CORRECT

['SENTENCE', 795405, ['Richard', 'firm'], 'LOW', 'CASE 1:', 0, [], 'CASE 2:', 0, [], 'CASE 3:', 1, 'CASE 4:', 1, ['Richard'], 'RULE:', 4]  
INCORRECT

['SENTENCE', 787143, ['Oceanographic', 'group'], 'LOW', 'CASE 1:', 0, [], 'CASE 2:', 0, [], 'CASE 3:', 1, 'CASE 4:', 1, ['Oceanographic'], 'RULE:', 4]  
INCORRECT

['SENTENCE', 753054, ['video-capture', 'DVD'], 'LOW', 'CASE 1:', 0, [], 'CASE 2:', 0, [], 'CASE 3:', 1, 'CASE 4:', 1, ['video-capture'], 'RULE:', 4]  
CORRECT

['SENTENCE', 750300, ['Riverkeeper', 'conservation'], 'LOW', 'CASE 1:', 0, [], 'CASE 2:', 0, [], 'CASE 3:', 1, 'CASE 4:', 1, ['Riverkeeper'], 'RULE:', 4]  
INCORRECT

['SENTENCE', 722637, ['white-water', 'trip'], 'LOW', 'CASE 1:', 0, [], 'CASE 2:', 0, [], 'CASE 3:', 1, 'CASE 4:', 1, ['white-water'], 'RULE:', 4]

['SENTENCE', 140156, ['child', 'others'], 'MEDIUM', 'CASE 1:', 0, [], 'CASE 2:', 0, [], 'CASE 3:', 1, 'CASE 4:', 1, ['others'], 'RULE:', 5]  
UNCERTAIN

['SENTENCE', 761491, ['megachurches', 'church'], 'LOW', 'CASE 1:', 0, [], 'CASE 2:', 0, [], 'CASE 3:', 1, 'CASE 4:', 1, ['megachurches'], 'RULE:', 5]  
CORRECT

['SENTENCE', 607299, ['ditch', 'anything'], 'LOW', 'CASE 1:', 0, [], 'CASE 2:', 0, [], 'CASE 3:', 1, 'CASE 4:', 1, ['anything'], 'RULE:', 5]  
INCORRECT

['SENTENCE', 567476, ['deconstructionists', 'theorist'], 'LOW', 'CASE 1:', 0, [], 'CASE 2:', 0, [], 'CASE 3:', 1, 'CASE 4:', 1, ['deconstructionists'], 'RULE:', 5]  
CORRECT

['SENTENCE', 435232, ['leather', 'shearling'], 'LOW', 'CASE 1:', 0, [], 'CASE 2:', 0, [], 'CASE 3:', 1, 'CASE 4:', 1, ['shearling'], 'RULE:', 5]  
UNCERTAIN

['SENTENCE', 189629, ['Diana', 'everybody'], 'LOW', 'CASE 1:', 0, [], 'CASE 2:', 0, [], 'CASE 3:', 1, 'CASE 4:', 1, ['everybody'], 'RULE:', 5]  
UNCERTAIN

['SENTENCE', 676199, [""in", 'strike'], 'LOW', 'CASE 1:', 0, [], 'CASE 2:', 0, [], 'CASE 3:', 1, 'CASE 4:', 1, [""in"], 'RULE:', 5]  
INCORRECT

['SENTENCE', 680284, ['Marvin', 'teammate'], 'LOW', 'CASE 1:', 0, [], 'CASE 2:', 0, [], 'CASE 3:', 1, 'CASE 4:', 1, ['Marvin'], 'RULE:', 5]  
INCORRECT

['SENTENCE', 728360, ['double-headers', 'Met'], 'LOW', 'CASE 1:', 0, [], 'CASE 2:', 0, [], 'CASE 3:', 1, 'CASE 4:', 1, ['double-headers'], 'RULE:', 5]  
INCORRECT

['SENTENCE', 744567, ['extra-small', 'graf'], 'LOW', 'CASE 1:', 0, [], 'CASE 2:', 0, [], 'CASE 3:', 1, 'CASE 4:', 1, ['extra-small'], 'RULE:', 5]  
UNCERTAIN

REPORT

No additional patterns have been used other than the DefaultNpPattern provided in the assignment.

Successes of the program include having pairs that are appropriately categorized in terms of the four cases. The regular expressions do implement the Hearst Patterns as asked in the assignment.

In general, sentences that are malformed and do not have a proper NP in them are ignored as expected even if a possible Hearst Pattern could be detected. Also there are no chunks that are produced when there is a verb phrase that comes after keywords of each Hearst Pattern. This removes sentences where there is no hyponym/hypernym relationship.

Shortcomings of the program include just using the DefaultNpPattern which doesn't include any other corner cases outside the regular expression such as a punctuation mark. Another shortcoming includes the extraction of hyponym/hypernym pairs. Extraction is done based on looking for any Part-Of-Speech tags that contain N in which the tagset identifies them as nouns. This leads to compound words being separate nouns as their own when forming pairs. This then can possibly lead to false negatives for case 3. Another shortcoming is the exact string of nouns may contain punctuation marks based on the sentence and so when finding synsets the result most of the time results in a blank set in which false positives are detected when looking through the results in the case 4 list. The use of lemmatization has both pros and cons since the results of the synsets are different from each other. Another slight shortcoming is the results of the chunks produced by the third Hearst pattern. The current approach is to find the word other tagged as an adjective. By this assumption results are mixed. Correct matches do show up however any other sentence without a hyponym/hypernym relationship will be matched by this pattern. Another shortcoming is that any abbreviations are separated as their own nouns. This leads to multiple False positives. With these shortcomings however they do come with the benefits of finding hyponym/hypernym relationships that have adjectives and so that case is covered. The current implementation of Hearst's patterns is very basic and is the case because any addition in my opinion would prove to be redundant. A recurring example is when there is a NP that has additional information followed by the regular pattern that comes after the keywords (such as, such <NPi ... NPn>) In this case the sentence does not get chunked because when looking for synsets with a noun with additional information proved to be useful and so extracting just the noun is effective enough. Another decision alongside this is to leave all the nouns as is. There was an idea to combine words and to use the WordNet compounds however that would require an additional layer of abstraction determine which nouns should combine together based on the context of the sentence. Another shortcoming of the program is the detection of rules 1 and 2 for the chunks. The program uses the index function and it only returns the first instance of the keywords (such,as) A future improvement would be to find all keywords and take those as the correct ones when extracting the nouns. This however does come with the assumption that keywords furthest in the sentence is the correct pattern Hearst will match to. Finally the last shortcoming of this program is that it cannot run through the entire corpus.



Current results are using the first 800000 sentences from nyt\_big. An implementation with dictionaries and other strategies to reduce overhead could possibly run better and possibly the entire corpus. Although another alternative is to run the entire corpus through partitions and then parse them all to find the actual results as if the program had successfully run through the entire corpus.