In [1]:

```
pip install mrakun
```

```
Collecting mrakun
  Downloading mrakun-0.44.tar.gz (4.2 MB)
Requirement already satisfied: nltk in c:\users\kruti\anaconda3\lib\site-packages
(from mrakun) (3.4.5)
Requirement already satisfied: networkx in c:\users\kruti\anaconda3\lib\site-
packages (from mrakun) (2.4)
Collecting editdistance
  Downloading editdistance-0.5.3-cp37-cp37m-win_amd64.whl (23 kB)
Requirement already satisfied: pandas in c:\users\kruti\anaconda3\lib\site-packages
(from mrakun) (1.0.1)
Requirement already satisfied: numpy in c:\users\kruti\anaconda3\lib\site-packages
(from mrakun) (1.18.1)
Collecting py3plex
  Downloading py3plex-0.83.tar.gz (151 kB)
Requirement already satisfied: six in c:\users\kruti\anaconda3\lib\site-packages
(from nltk->mrakun) (1.14.0)
Requirement already satisfied: decorator>=4.3.0 in c:\users\kruti\anaconda3\lib\
site-packages (from networkx->mrakun) (4.4.1)
Requirement already satisfied: python-dateutil>=2.6.1 in c:\users\kruti\anaconda3\
lib\site-packages (from pandas->mrakun) (2.8.1)
Requirement already satisfied: pytz>=2017.2 in c:\users\kruti\anaconda3\lib\site-
packages (from pandas->mrakun) (2019.3)
Collecting rdflib>=0.1
  Downloading rdflib-5.0.0-py3-none-any.whl (231 kB)
Requirement already satisfied: scipy>=1.1.0 in c:\users\kruti\anaconda3\lib\site-
packages (from py3plex->mrakun) (1.4.1)
Collecting plotnine
  Downloading plotnine-0.7.1-py3-none-any.whl (4.4 MB)
Requirement already satisfied: cython>=0.20 in c:\users\kruti\anaconda3\lib\site-
packages (from py3plex->mrakun) (0.29.14)
Requirement already satisfied: tqdm>0.0 in c:\users\kruti\anaconda3\lib\site-
packages (from py3plex->mrakun) (4.42.1)
Requirement already satisfied: matplotlib>=3.0 in c:\users\kruti\anaconda3\lib\site-
packages (from py3plex->mrakun) (3.1.3)
Requirement already satisfied: gensim in c:\users\kruti\anaconda3\lib\site-packages
(from py3plex->mrakun) (3.8.3)
Requirement already satisfied: scikit-learn in c:\users\kruti\anaconda3\lib\site-
packages (from py3plex->mrakun) (0.22.1)
Requirement already satisfied: bitarray==1.2.1 in c:\users\kruti\anaconda3\lib\site-
packages (from py3plex->mrakun) (1.2.1)
Requirement already satisfied: seaborn in c:\users\kruti\anaconda3\lib\site-packages
(from py3plex->mrakun) (0.10.0)
Collecting isodate
```

```
  Downloading isodate-0.6.0-py2.py3-none-any.whl (45 kB)
Requirement already satisfied: pyparsing in c:\users\kruti\anaconda3\lib\site-
packages (from rdflib>=0.1->py3plex->mrakun) (2.4.6)
Collecting descartes>=1.1.0
  Downloading descartes-1.1.0-py3-none-any.whl (5.8 kB)
Collecting mizani>=0.7.1
  Downloading mizani-0.7.2-py3-none-any.whl (62 kB)
Requirement already satisfied: patsy>=0.5.1 in c:\users\kruti\anaconda3\lib\site-
packages (from plotnine->py3plex->mrakun) (0.5.1)
Collecting statsmodels>=0.11.1
  Downloading statsmodels-0.12.1-cp37-none-win_amd64.whl (9.1 MB)
Requirement already satisfied: kiwisolver>=1.0.1 in c:\users\kruti\anaconda3\lib\
site-packages (from matplotlib>=3.0->py3plex->mrakun) (1.1.0)
Requirement already satisfied: cycler>=0.10 in c:\users\kruti\anaconda3\lib\site-
packages (from matplotlib>=3.0->py3plex->mrakun) (0.10.0)
Requirement already satisfied: smart-open>=1.8.1 in c:\users\kruti\anaconda3\lib\
site-packages (from gensim->py3plex->mrakun) (2.0.0)
Requirement already satisfied: joblib>=0.11 in c:\users\kruti\anaconda3\lib\site-
packages (from scikit-learn->py3plex->mrakun) (0.14.1)
Collecting palettable
  Downloading palettable-3.3.0-py2.py3-none-any.whl (111 kB)
Requirement already satisfied: setuptools in c:\users\kruti\anaconda3\lib\site-
packages (from kiwisolver>=1.0.1->matplotlib>=3.0->py3plex->mrakun)
(45.2.0.post20200210)
Requirement already satisfied: requests in c:\users\kruti\anaconda3\lib\site-
packages (from smart-open>=1.8.1->gensim->py3plex->mrakun) (2.22.0)
Requirement already satisfied: boto3 in c:\users\kruti\anaconda3\lib\site-packages
(from smart-open>=1.8.1->gensim->py3plex->mrakun) (1.14.8)
Requirement already satisfied: boto in c:\users\kruti\anaconda3\lib\site-packages
(from smart-open>=1.8.1->gensim->py3plex->mrakun) (2.49.0)
Requirement already satisfied: urllib3!=1.25.0,!=1.25.1,<1.26,>=1.21.1 in c:\users\
kruti\anaconda3\lib\site-packages (from requests->smart-open>=1.8.1->gensim-
>py3plex->mrakun) (1.25.8)
Requirement already satisfied: idna<2.9,>=2.5 in c:\users\kruti\anaconda3\lib\site-
packages (from requests->smart-open>=1.8.1->gensim->py3plex->mrakun) (2.8)
Requirement already satisfied: certifi>=2017.4.17 in c:\users\kruti\anaconda3\lib\
site-packages (from requests->smart-open>=1.8.1->gensim->py3plex->mrakun)
(2019.11.28)
Requirement already satisfied: chardet<3.1.0,>=3.0.2 in c:\users\kruti\anaconda3\
lib\site-packages (from requests->smart-open>=1.8.1->gensim->py3plex->mrakun)
(3.0.4)
Requirement already satisfied: s3transfer<0.4.0,>=0.3.0 in c:\users\kruti\anaconda3\
lib\site-packages (from boto3->smart-open>=1.8.1->gensim->py3plex->mrakun) (0.3.3)
Requirement already satisfied: jmespath<1.0.0,>=0.7.1 in c:\users\kruti\anaconda3\
lib\site-packages (from boto3->smart-open>=1.8.1->gensim->py3plex->mrakun) (0.10.0)
Requirement already satisfied: botocore<1.18.0,>=1.17.8 in c:\users\kruti\anaconda3\
lib\site-packages (from boto3->smart-open>=1.8.1->gensim->py3plex->mrakun) (1.17.8)
```

```
Requirement already satisfied: docutils<0.16,>=0.10 in c:\users\kruti\anaconda3\lib\
site-packages (from botocore<1.18.0,>=1.17.8->boto3->smart-open>=1.8.1->gensim-
>py3plex->mrakun) (0.15.2)
Building wheels for collected packages: mrakun, py3plex
  Building wheel for mrakun (setup.py): started
  Building wheel for mrakun (setup.py): finished with status 'done'
  Created wheel for mrakun: filename=mrakun-0.44-py3-none-any.whl size=18984
sha256=ff9081bcc73da409c942a933b2bbdd0a6fcdc4a8f0356596e0bdef294b026283
  Stored in directory: c:\users\kruti\appdata\local\pip\cache\wheels\1d\
78\11\24c1910306acf3c00d7e3ec5423453347aa5e8a1c7521305e1
  Building wheel for py3plex (setup.py): started
  Building wheel for py3plex (setup.py): finished with status 'done'
  Created wheel for py3plex: filename=py3plex-0.83-py3-none-any.whl size=182954
sha256=3754a69bad1f6be731832de866e8d246505a61ca0bebffc46f8b329828062af9
  Stored in directory: c:\users\kruti\appdata\local\pip\cache\wheels\08\fe\cd\
284395dbd200a614679f228818d0c41f2c053132edb464045e
Successfully built mrakun py3plex
Installing collected packages: editdistance, isodate, rdflib, descartes, palettable,
mizani, statsmodels, plotnine, py3plex, mrakun
  Attempting uninstall: statsmodels
    Found existing installation: statsmodels 0.11.0
    Uninstalling statsmodels-0.11.0:
      Successfully uninstalled statsmodels-0.11.0
Successfully installed descartes-1.1.0 editdistance-0.5.3 isodate-0.6.0 mizani-0.7.2
mrakun-0.44 palettable-3.3.0 plotnine-0.7.1 py3plex-0.83 rdflib-5.0.0 statsmodels-
0.12.1
Note: you may need to restart the kernel to use updated packages.

ERROR: mizani 0.7.2 has requirement pandas>=1.1.0, but you'll have pandas 1.0.1
which is incompatible.
ERROR: plotnine 0.7.1 has requirement pandas>=1.1.0, but you'll have pandas 1.0.1
which is incompatible.
```

In [2]:

```
from mrakun import RakunDetector
import nltk
# nltk.download('punkt')
import nltk
# nltk.download('stopwords')
from nltk.corpus import stopwords

blob_of_text = "Brexit (/'brɛksɪt, 'brɛgzɪt/;[1] a portmanteau of \"British\"
and \"exit\") is the scheduled withdrawal of the United Kingdom (UK) from the
European Union (EU). Following a June 2016 referendum, in which 51.9% voted to
leave, the UK government formally announced the country's withdrawal in March 2017,
starting a two-year process that was due to conclude with the UK withdrawing on 29
March 2019. As the UK parliament thrice voted against the negotiated withdrawal
```

agreement, that deadline has been extended twice, and is currently 31 October 2019.[2][3] An Act of Parliament requires the government to seek a third extension if no agreement is reached before 19 October. Withdrawal is advocated by Eurosceptics and opposed by pro-Europeanists, both of whom span the political spectrum. The UK joined the European Communities (EC) in 1973, with continued membership endorsed in a 1975 referendum. In the 1970s and 1980s, withdrawal from the EC was advocated mainly by the political left, e.g. in the Labour Party's 1983 election manifesto. From the 1990s, the eurosceptic wing of the Conservative Party grew, and led a rebellion over ratification of the 1992 Maastricht Treaty that established the EU. In parallel with the UK Independence Party (UKIP), and the cross-party People's Pledge campaign, it pressured Conservative Prime Minister David Cameron to hold a referendum on continued EU membership. Cameron, who had campaigned to remain, resigned after the result and was succeeded by Theresa May. On 29 March 2017, the UK government invoked Article 50 of the Treaty on European Union, formally starting the withdrawal. May called a snap general election in June 2017, which resulted in a Conservative minority government supported by the Democratic Unionist Party. UK–EU withdrawal negotiations began later that month. The UK negotiated to leave the EU customs union and single market. This resulted in the November 2018 withdrawal agreement, but the UK parliament voted against ratifying it three times. The Labour Party wanted any agreement to maintain a customs union, while many Conservatives opposed the agreement's financial settlement on the UK's share of EU financial obligations, as well as the Irish backstop designed to prevent border controls in Ireland. The Liberal Democrats, Scottish National Party and others seek to reverse Brexit through a second referendum. The EU has declined a re-negotiation that omits the backstop. In March 2019, the UK parliament voted for May to ask the EU to delay Brexit until October. Having failed to pass her agreement, May resigned as Prime Minister in July and was succeeded by Boris Johnson. He sought to replace parts of the agreement and vowed to leave the EU by the new deadline, with or without an agreement."

```
hyperparameters = {"distance_threshold":2,
                   "distance_method": "editdistance",
                   "num_keywords" : 10,
                   "pair_diff_length":2,
                   "stopwords" : stopwords.words('english'),
                   "bigram_count_threshold":2,
                   "num_tokens":[1,2],
                   "max_similar" : 3, ## n most similar can show up n times
                   "max_occurrence" : 3} ## maximum frequency overall

keyword_detector = RakunDetector(hyperparameters)
keywords = keyword_detector.find_keywords(blob_of_text, input_type = "text")
print(keywords)
keyword_detector.visualize_network()
```

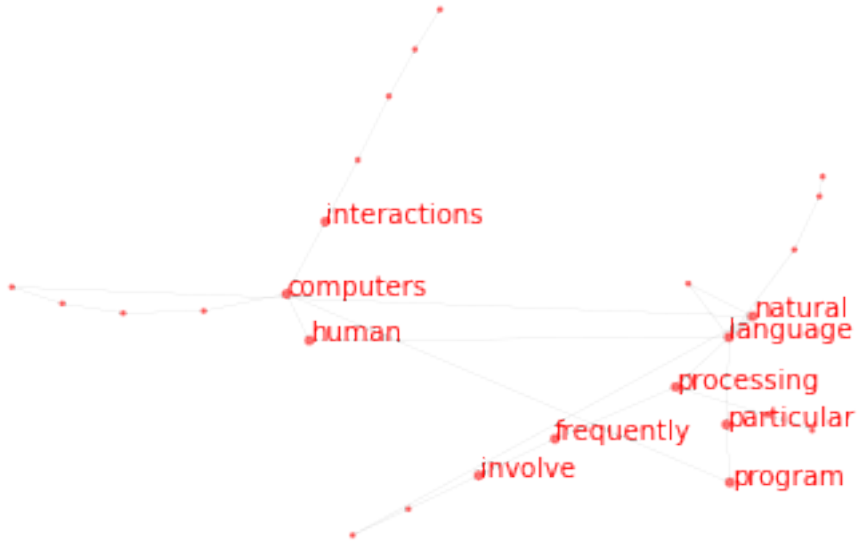
```
25-Nov-20 05:47:35 - Initiated a keyword detector instance.
25-Nov-20 05:47:35 - Number of nodes reduced from 136 to 129
25-Nov-20 05:47:35 - Name:
Type: DiGraph
```

```
Number of nodes: 129
Number of edges: 182
Average in degree:   1.4109
Average out degree:   1.4109
  8%|
| 4/50 [00:00<00:01, 24.56it/s]

[('withdrawal', 0.36376312335958005), ('party', 0.2600424458661417), ('agreement',
0.24967960547900261), ('march', 0.20897924868766404), ('voted',
0.17492002952755906), ('starting', 0.1384873277559055), ('government',
0.1373185285433071), ('referendum', 0.12309301181102363), ('parliament',
0.11887918307086615), ('union', 0.11731053149606299)]
Beginning parsing..
Using custom init positions!

100%|
|
50/50 [00:02<00:00, 22.40it/s]

BarnesHut Approximation  took  0.27  seconds
Repulsion forces  took  1.75  seconds
Gravitational forces  took  0.02  seconds
Attraction forces  took  0.07  seconds
AdjustSpeedAndApplyForces step  took  0.08  seconds
```

In [3]:

```
blob_of_text = "Natural language processing (NLP) is a subfield of linguistics,
computer science, and artificial intelligence concerned with the interactions
between computers and human language, in particular how to program computers to
process and analyze large amounts of natural language data.Challenges in natural
language processing frequently involve speech recognition, natural language
understanding, and natural-language generation."
hyperparameters = {"distance_threshold":2,
```

```
                    "distance_method": "editdistance",
                    "num_keywords" : 10,
                    "pair_diff_length":2,
                    "stopwords" : stopwords.words('english'),
                    "bigram_count_threshold":2,
                    "num_tokens":[1,2],
                    "max_similar" : 3, ## n most similar can show up n times
                    "max_occurrence" : 3} ## maximum frequency overall

keyword_detector = RakunDetector(hyperparameters)
keywords = keyword_detector.find_keywords(blob_of_text, input_type = "text")
print(keywords)
keyword_detector.visualize_network()

25-Nov-20 05:47:50 - Initiated a keyword detector instance.
25-Nov-20 05:47:50 - Number of nodes reduced from 27 to 26
25-Nov-20 05:47:50 - Name:
Type: DiGraph
Number of nodes: 26
Number of edges: 29
Average in degree:   1.1154
Average out degree:   1.1154
 58%|████████████████████████████████████████
| 29/50 [00:00<00:00, 289.94it/s]

[('language', 0.515), ('natural language', 0.4008333333333334), ('computers',
0.31833333333333336), ('human', 0.195), ('processing', 0.18333333333333335),
('program', 0.14333333333333334), ('particular', 0.14333333333333334),
('interactions', 0.14), ('recognition', 0.11333333333333334), ('involve',
0.11333333333333334)]
Beginning parsing..
Using custom init positions!

100%|
████████████████████████████████████████████████████████████████████████████|
50/50 [00:00<00:00, 286.48it/s]

BarnesHut Approximation  took  0.03  seconds
Repulsion forces  took  0.11  seconds
Gravitational forces  took  0.00  seconds
Attraction forces  took  0.01  seconds
AdjustSpeedAndApplyForces step  took  0.02  seconds
```



In [4]:

```
from mrakun import RakunDetector
from nltk.corpus import stopwords

hyperparameters = {"distance_threshold":3,
                   "num_keywords" : 10,
                   "distance_method": "editdistance",
                   "pair_diff_length":2,
                   "stopwords" : stopwords.words('english'),
                   "bigram_count_threshold":2,
                   "num_tokens":[1]}

keyword_detector = RakunDetector(hyperparameters)
example_data = "NLP_sample.txt"
keywords = keyword_detector.find_keywords(example_data)
print(keywords)

keyword_detector.verbose = False
# KEYWORDS WITH THEIR CENTALITY SCORES ARE RETURNED
```

```
25-Nov-20 05:48:46 - Initiated a keyword detector instance.
25-Nov-20 05:48:47 - Number of nodes reduced from 334 to 299

[('language', 0.33229926972935875), ('processing', 0.2327731390464424), ('natural',
0.15974034717043678), ('machine', 0.15795915907772734), ('human',
0.12883872279845435), ('information', 0.12205952138838044), ('systems',
0.11825563615272788), ('translation', 0.11569837073192782), ('learning',
0.10850676790274105), ('input', 0.09889654373714776)]
```

In [5]:

```
keyword_detector.visualize_network()
```

4%|

```
| 2/50 [00:00<00:03, 15.08it/s]

Beginning parsing..
Using custom init positions!

100%|
```

```
50/50 [00:08<00:00,  5.68it/s]

BarnesHut Approximation  took  0.66  seconds
Repulsion forces  took  7.67  seconds
Gravitational forces  took  0.05  seconds
Attraction forces  took  0.14  seconds
AdjustSpeedAndApplyForces step  took  0.19  seconds
```