

Question 1: Create lexical chains based on the lexical relationships of synonymy, antonymy, and one level of hyper/hyponymy. In other words, if a noun is related by one of these relationships to an existing lexical chain, it should be added to it, otherwise it should start a new chain.

Answer:

Algorithm:

1. Extract all the nouns ignoring proper nouns from the input text.
2. For each noun in the text
 - a. Find synonyms, antonyms, hypernyms and hyponyms using WordNet with NLTK library.
 - b. If the noun matches any of the noun in the existing lexical chains, then increment the word_frequency for that noun by 1.
 - c. Else If the synonyms or antonyms or hypernyms or hyponyms match any existing nouns in any of the lexical chains; add the noun to that lexical chain. If not create a new lexical chain and add the noun with word_frequency 1.
 - d. Print all the lexical Chains

Source code: lexical_chains_for_summary.py

Input Text File: input_text.txt

Input Text:

Labour plans maternity pay rise. Maternity pay for new mothers is to rise by £1,400 as part of new proposals announced by the Trade and Industry Secretary Patricia Hewitt. It would mean paid leave would be increased to nine months by 2007, Ms Hewitt told GMTV's Sunday programme. Other plans include letting maternity pay be given to fathers and extending rights to parents of older children. The Tories dismissed the maternity pay plan as "desperate", while the Liberal Democrats said it was misdirected. Ms Hewitt said: "We have already doubled the length of maternity pay, it was 13 weeks when we were elected, we have already taken it up to 26 weeks. "We are going to extend the pay to nine months by 2007 and the aim is to get it right up to the full 12 months by the end of the next Parliament." She said new mothers were already entitled to 12 months leave, but that many women could not take it as only six of those months were paid. "We have made a firm commitment. We will definitely extend the maternity pay, from the six months where it now is to nine months, that's the extra £1,400." She said ministers would consult on other proposals that could see fathers being allowed to take some of their partner's maternity pay or leave period, or extending the rights of flexible working to carers or parents of older children. The Shadow Secretary of State for the Family, Theresa May, said: "These plans were announced by Gordon Brown in his pre-budget review in December and Tony Blair is now recycling it in his desperate bid to win back women voters." She said the Conservatives would announce their proposals closer to the General Election. Liberal Democrat spokeswoman for women Sandra Gidley said: "While mothers would welcome any extra maternity pay the Liberal Democrats feel this money is being misdirected." She said her party would boost maternity pay in the first six months to allow more women to stay at home in that time. Ms Hewitt also stressed

the plans would be paid for by taxpayers, not employers. But David Frost, director general of the British Chambers of Commerce, warned that many small firms could be "crippled" by the move. "While the majority of any salary costs may be covered by the government's statutory pay, recruitment costs, advertising costs, retraining costs and the strain on the company will not be," he said. Further details of the government's plans will be outlined on Monday. New mothers are currently entitled to 90% of average earnings for the first six weeks after giving birth, followed by £102.80 a week until the baby is six months old.

List of nouns in the text::

['maternity', 'pay', 'rise', 'pay', 'mothers', '£1,400', 'part', 'proposals', 'leave', 'months', 'programme', 'plans', 'maternity', 'pay', 'fathers', 'rights', 'parents', 'children', 'maternity', 'pay', 'plan', 'length', 'maternity', 'pay', 'weeks', 'weeks', 'pay', 'months', 'aim', 'months', 'end', 'mothers', 'months', 'women', 'months', 'commitment', 'maternity', 'pay', 'months', 'months', '£1,400', 'ministers', 'proposals', 'fathers', 'partner', 'maternity', 'pay', 'period', 'rights', 'carers', 'parents', 'children', 'plans', 'pre-budget', 'review', 'bid', 'women', 'voters', 'proposals', 'spokeswoman', 'women', 'mothers', 'maternity', 'money', 'party', 'maternity', 'pay', 'months', 'women', 'home', 'time', 'plans', 'taxpayers', 'employers', 'director', 'general', 'Chambers', 'firms', 'move', 'majority', 'costs', 'government', 'pay', 'recruitment', 'costs', 'advertising', 'costs', 'costs', 'strain', 'company', 'details', 'government', 'plans', 'mothers', '%', 'earnings', 'weeks', 'birth', 'week', 'baby', 'months']

All lexical chains:

[{'maternity': 8}, {'pay': 10, 'earnings': 1}, {'rise': 1, 'move': 1}, {'mothers': 4}, {'£1,400': 2}, {'part': 1, 'leave': 1, 'end': 1, 'period': 1, 'months': 1, 'time': 1, 'move': 1, 'strain': 1, 'details': 1, 'weeks': 1, 'week': 1}, {'proposals': 3}, {'months': 9}, {'programme': 1, 'plans': 1, 'plan': 1, 'aim': 1, 'end': 1, 'period': 1, 'months': 1, 'time': 1, 'weeks': 1, 'week': 1}, {'fathers': 2}, {'rights': 2}, {'parents': 2}, {'children': 2}, {'length': 1}, {'weeks': 3}, {'women': 4}, {'commitment': 1}, {'ministers': 1}, {'partner': 1, 'move': 1}, {'carers': 1}, {'pre-budget': 1}, {'review': 1}, {'bid': 1}, {'voters': 1}, {'spokeswoman': 1}, {'money': 1}, {'party': 1, 'company': 1}, {'home': 1}, {'taxpayers': 1}, {'employers': 1}, {'director': 1}, {'general': 1}, {'Chambers': 1}, {'firms': 1}, {'majority': 1}, {'costs': 4}, {'government': 2}, {'recruitment': 1}, {'advertising': 1}, {'%': 1}, {'birth': 1}, {'baby': 1}]

Lexical chains in decreasing order of their lengths:

[{'part': 1, 'leave': 1, 'end': 1, 'period': 1, 'months': 1, 'time': 1, 'move': 1, 'strain': 1, 'details': 1, 'weeks': 1, 'week': 1}, {'programme': 1, 'plans': 1, 'plan': 1, 'aim': 1, 'end': 1, 'period': 1, 'months': 1, 'time': 1, 'weeks': 1, 'week': 1}, {'pay': 10, 'earnings': 1}, {'rise': 1, 'move': 1}, {'partner': 1, 'move': 1}, {'party': 1, 'company': 1}, {'maternity': 8}, {'mothers': 4}, {'£1,400': 2}, {'proposals': 3}, {'months': 9}, {'fathers': 2}, {'rights': 2}, {'parents': 2}, {'children': 2}, {'length': 1}, {'weeks': 3}, {'women': 4}, {'commitment': 1}, {'ministers': 1}, {'carers': 1}, {'pre-budget': 1}, {'review': 1}, {'bid': 1}, {'voters': 1}, {'spokeswoman': 1}, {'money': 1}, {'home': 1}, {'taxpayers': 1}, {'employers': 1}, {'director': 1}, {'general': 1}, {'Chambers': 1}, {'firms': 1}, {'majority': 1}, {'costs': 4}, {'government': 2}, {'recruitment': 1}, {'advertising': 1}, {'%': 1}, {'birth': 1}, {'baby': 1}]

Question 2: Evaluate the results (qualitatively) and comment on how they might be improved.

Answer: All lexical chains:

[{'maternity': 8}, {'pay': 10, 'earnings': 1}, {'rise': 1, 'move': 1}, {'mothers': 4}, {'£1,400': 2}, {'part': 1, 'leave': 1, 'end': 1, 'period': 1, 'months': 1, 'time': 1, 'move': 1, 'strain': 1, 'details': 1, 'weeks': 1, 'week': 1}, {'proposals': 3}, {'months': 9}, {'programme': 1, 'plans': 1, 'plan': 1, 'aim': 1, 'end': 1, 'period': 1, 'months': 1, 'time': 1, 'weeks': 1, 'week': 1}, {'fathers': 2}, {'rights': 2}, {'parents': 2}, {'children': 2}, {'length': 1}, {'weeks': 3}, {'women': 4}, {'commitment': 1}, {'ministers': 1}, {'partner': 1, 'move': 1}, {'carers': 1}, {'pre-budget': 1}, {'review': 1}, {'bid': 1}, {'voters': 1}, {'spokeswoman': 1}, {'money': 1}, {'party': 1, 'company': 1}, {'home': 1}, {'taxpayers': 1}, {'employers': 1}, {'director': 1}, {'general': 1}, {'Chambers': 1}, {'firms': 1}, {'majority': 1}, {'costs': 4}, {'government': 2}, {'recruitment': 1}, {'advertising': 1}, {'%': 1}, {'birth': 1}, {'baby': 1}]

Lexical chains in decreasing order of their lengths:

[{'part': 1, 'leave': 1, 'end': 1, 'period': 1, 'months': 1, 'time': 1, 'move': 1, 'strain': 1, 'details': 1, 'weeks': 1, 'week': 1}, {'programme': 1, 'plans': 1, 'plan': 1, 'aim': 1, 'end': 1, 'period': 1, 'months': 1, 'time': 1, 'weeks': 1, 'week': 1}, {'pay': 10, 'earnings': 1}, {'rise': 1, 'move': 1}, {'partner': 1, 'move': 1}, {'party': 1, 'company': 1}, {'maternity': 8}, {'mothers': 4}, {'£1,400': 2}, {'proposals': 3}, {'months': 9}, {'fathers': 2}, {'rights': 2}, {'parents': 2}, {'children': 2}, {'length': 1}, {'weeks': 3}, {'women': 4}, {'commitment': 1}, {'ministers': 1}, {'carers': 1}, {'pre-budget': 1}, {'review': 1}, {'bid': 1}, {'voters': 1}, {'spokeswoman': 1}, {'money': 1}, {'home': 1}, {'taxpayers': 1}, {'employers': 1}, {'director': 1}, {'general': 1}, {'Chambers': 1}, {'firms': 1}, {'majority': 1}, {'costs': 4}, {'government': 2}, {'recruitment': 1}, {'advertising': 1}, {'%': 1}, {'birth': 1}, {'baby': 1}]

Evaluation:

1. Some of the related words are not in the same lexical chain because we are only using one level of hypernyms or hyponyms.
2. Pronouns are being ignored.
3. The criteria for noun to be added to a particular lexical chain can be more refined.

Improvements:

1. If the algorithm comes across a word that belongs to multiple lexical chains it can use following measures to decide which lexical chain should the word be added:
 - a. If it has relations to words the algorithm can look for recency - which chain is it closer to.
 - b. To which chain does it has relationship to more words
2. To add pronouns to lexical chains check for saliency - being part of a longer chain.
3. Use sibling hypo/hypernymy and meronym relationships to form better lexical chains - words related to each other would get added to one lexical chain.
4. Use phrasal nouns by finding short sequences of word from wordnet.
5. Use a threshold number of words between the noun and the closest element of a chain as a criteria to add a noun to a particular lexical chain.
6. Restrict the senses of the nouns in the chain to those consistent with all nouns in the chain - just those senses related by the given lexical relations.

Question 4: Extra credit: Use the lexical chains to automatically create a summary of the input article.

Answer:

Algorithm:

1. Find strong chains from the list of lexical chains formed in question 1. We filter out the strong chains based on the length of the chains or sum of word frequencies in that chain.
2. Weight each sentence in the input text using strong lexical chains. The weight of a sentence is calculated by the sum of word frequencies (from the strong lexical chains) of all the words that occur in any of strong chain.
3. Sort the sentences in descending order of their weights calculated.
4. Calculate the length of summary text as $\frac{1}{3}$ of the input text.
5. Add the sentences (number of sentences to be picked = length of summary text) from the descending order sorted list into summary text in the order in which they appear in input text.
6. Print the summary text.

Source code: lexical_chains_for_summary.py

Input Text File: input_text.txt

All lexical chains::

```
[{'maternity': 8}, {'pay': 10, 'earnings': 1}, {'rise': 1, 'move': 1}, {'mothers': 4}, {'£1,400': 2}, {'part': 1, 'leave': 1, 'end': 1, 'period': 1, 'months': 1, 'time': 1, 'move': 1, 'strain': 1, 'details': 1, 'weeks': 1, 'week': 1}, {'proposals': 3}, {'months': 9}, {'programme': 1, 'plans': 1, 'plan': 1, 'aim': 1, 'end': 1, 'period': 1, 'months': 1, 'time': 1, 'weeks': 1, 'week': 1}, {'fathers': 2}, {'rights': 2}, {'parents': 2}, {'children': 2}, {'length': 1}, {'weeks': 3}, {'women': 4}, {'commitment': 1}, {'ministers': 1}, {'partner': 1, 'move': 1}, {'carers': 1}, {'pre-budget': 1}, {'review': 1}, {'bid': 1}, {'voters': 1}, {'spokeswoman': 1}, {'money': 1}, {'party': 1, 'company': 1}, {'home': 1}, {'taxpayers': 1}, {'employers': 1}, {'director': 1}, {'general': 1}, {'Chambers': 1}, {'firms': 1}, {'majority': 1}, {'costs': 4}, {'government': 2}, {'recruitment': 1}, {'advertising': 1}, {'%': 1}, {'birth': 1}, {'baby': 1}]
```

Lexical chains in decreasing order of their lengths:

```
[{'part': 1, 'leave': 1, 'end': 1, 'period': 1, 'months': 1, 'time': 1, 'move': 1, 'strain': 1, 'details': 1, 'weeks': 1, 'week': 1}, {'programme': 1, 'plans': 1, 'plan': 1, 'aim': 1, 'end': 1, 'period': 1, 'months': 1, 'time': 1, 'weeks': 1, 'week': 1}, {'pay': 10, 'earnings': 1}, {'rise': 1, 'move': 1}, {'partner': 1, 'move': 1}, {'party': 1, 'company': 1}, {'maternity': 8}, {'mothers': 4}, {'£1,400': 2}, {'proposals': 3}, {'months': 9}, {'fathers': 2}, {'rights': 2}, {'parents': 2}, {'children': 2}, {'length': 1}, {'weeks': 3}, {'women': 4}, {'commitment': 1}, {'ministers': 1}, {'carers': 1}, {'pre-budget': 1}, {'review': 1}, {'bid': 1}, {'voters': 1}, {'spokeswoman': 1}, {'money': 1}, {'home': 1}, {'taxpayers': 1}, {'employers': 1}, {'director': 1}, {'general': 1}, {'Chambers': 1}, {'firms': 1}, {'majority': 1}, {'costs': 4}, {'government': 2}, {'recruitment': 1}, {'advertising': 1}, {'%': 1}, {'birth': 1}, {'baby': 1}]
```

Strong chain final list:

[{'part': 1, 'leave': 1, 'end': 1, 'period': 1, 'months': 1, 'time': 1, 'move': 1, 'strain': 1, 'details': 1, 'weeks': 1, 'week': 1}, {'pay': 10, 'earnings': 1}]

Sentence score dictionary:

{'Labour plans maternity pay rise.': 10, 'Maternity pay for new mothers is to rise by £1,400 as part of new proposals announced by the Trade and Industry Secretary Patricia Hewitt.': 11, 'It would mean paid leave would be increased to nine months by 2007, Ms Hewitt told GMTV's Sunday programme.': 2, 'Other plans include letting maternity pay be given to fathers and extending rights to parents of older children.': 10, 'The Tories dismissed the maternity pay plan as "desperate", while the Liberal Democrats said it was misdirected.': 10, 'Ms Hewitt said: "We have already doubled the length of maternity pay, it was 13 weeks when we were elected, we have already taken it up to 26 weeks.': 12, "'We are going to extend the pay to nine months by 2007 and the aim is to get it right up to the full 12 months by the end of the next Parliament."': 13, 'She said new mothers were already entitled to 12 months leave, but that many women could not take it as only six of those months were paid.': 3, "'We have made a firm commitment.': 0, 'We will definitely extend the maternity pay, from the six months where it now is to nine months, that's the extra £1,400.'": 12, "She said ministers would consult on other proposals that could see fathers being allowed to take some of their partner's maternity pay or leave period, or extending the rights of flexible working to carers or parents of older children."': 12, 'The Shadow Secretary of State for the Family, Theresa May, said: "These plans were announced by Gordon Brown in his pre-budget review in December and Tony Blair is now recycling it in his desperate bid to win back women voters."': 0, 'She said the Conservatives would announce their proposals closer to the General Election.': 0, 'Liberal Democrat spokeswoman for women Sandra Gidley said: "While mothers would welcome any extra maternity pay the Liberal Democrats feel this money is being misdirected."': 10, 'She said her party would boost maternity pay in the first six months to allow more women to stay at home in that time.': 12, 'Ms Hewitt also stressed the plans would be paid for by taxpayers, not employers.': 0, 'But David Frost, director general of the British Chambers of Commerce, warned that many small firms could be "crippled" by the move.': 1, "'While the majority of any salary costs may be covered by the government's statutory pay, recruitment costs, advertising costs, retraining costs and the strain on the company will not be," he said.': 11, "Further details of the government's plans will be outlined on Monday."': 1, 'New mothers are currently entitled to 90% of average earnings for the first six weeks after giving birth, followed by £102.80 a week until the baby is six months old.': 4}

Summary Length:

6

Summary:

Maternity pay for new mothers is to rise by £1,400 as part of new proposals announced by the Trade and Industry Secretary Patricia Hewitt. Ms Hewitt said: "We have already doubled the length of maternity pay, it was 13 weeks when we were elected, we have already taken it up to

26 weeks."We are going to extend the pay to nine months by 2007 and the aim is to get it right up to the full 12 months by the end of the next Parliament."We will definitely extend the maternity pay, from the six months where it now is to nine months, that's the extra £1,400."She said ministers would consult on other proposals that could see fathers being allowed to take some of their partner's maternity pay or leave period, or extending the rights of flexible working to carers or parents of older children.She said her party would boost maternity pay in the first six months to allow more women to stay at home in that time.