Type: MCQ

Q1. The knowledge needed to order and group words together is ____ (0.5)

1. **Syntax
2. Morphology
3. Phonology
4. Semantics

Q2. Which of the following step can be used in lexical disambiguation? (0.5)

1. **Part-of-Speech Tagging
2. Syntactic disambiguation
3. Discourse
4. Morphology

Q3. English mainly uses prefixes and suffixes to express _____ morphology. (0.5)

1. Only inflectional
2. Only derivational
3. **Both inflectional and derivational
4. Cliticization

Q4. Which of the following class is called productive (0.5)

1. **regular verbs
2. irregular verbs
3. irregular nouns
4. irregular verbs in past form

Q5. Which of the following is FALSE? (0.5)

1. FSA can only be used as a word recognizer
2. **FST can only be used as a word recognizer
3. FST can be used for as a word recognizer and generator
4. FST can be used as a relater

Q6. The limitation of Laplace smoothing is that (0.5)

1. It is same as Maximum Likelihood Estimation (MLE)
2. **Too much probability mass is moved to all the zeros

3. It can be applied only to bigrams model
4. It can be applied only to unigram model

Q7. Consider vocabulary as {a, b, c} and corpus as

b a b a a c b c a c a c

The set of unseen bigram/s is/are: (0.5)

1. **{bb, cc}
2. {b}
3. {bc, cc}
4. {ab, ac}

Q8. Given the bigram counts

	i	want	to	eat
i	5	827	0	9
want	2	0	608	1
to	2	0	4	686

and unigram counts

i	want	to
2533	927	2417

The probability P(want | i) is: (0.5)

1. **0.326
2. 0.002
3. 0.00355
4. 0.002

Q9. _____ method uses the frequency of singletons as a re-estimate of the frequency of zero-count bigrams. (0.5)

1. Add-1 smoothing
2. Laplace smoothing
3. Add-k smoothing
4. **Good-Turing

Q10. Minimum Edit Distance algorithm uses (0.5)

1. **dynamic programming
2. greedy strategy
3. brute force technique
4. divide and conquer

Type: DES

Q11. Fill the table below using Minimum Edit Distance algorithm

T	6	↓ 5	← ↙↓ 6	↙ 5	← 6	←↓ 7	← ↙↓ 8	↙↓ 7
N	5	↓ 4	←↓ 5	← ↙↓ 6	← ↙↓ 7	↙↓ 6	← ↙↓ 7	↓ 6
E	4	↙ 3	← 4	←↓ 5	← ↙↓ 6	↓ 5	← ↙↓ 6	↓ 5
T	3	← ↙↓ 4	← ↙↓ 5	↙ 4	← 5	↓ 4	← ↙↓ 5	↙ 4
N	2	← ↙↓ 3	← ↙↓ 4	← ↙↓ 5	← ↙↓ 6	↙ 3	← 4	← 5
I	1	← ↙↓ 2	← ↙↓ 3	← ↙↓ 4	↙ 3	← 4	← 5	← 6
#	0	1	2	3	4	5	6	7
	#	E	X	T	I	N	C	T

i) Find the Minimum Edit Distance and the operation list

Operation List: D-Delete, S-Substitute, and I-Insert

* * * I N * T E N T

E X T I N C T * * *

Operation List

I, I , I, S, S , I, S, D, D, D

ii) Show the backtrace and corresponding alignment (4)

Any one backtrace and the corresponding alignment out of three possibilities

Q12. Consider the following training data

<s> peter piper picked a peck of pickled peppers </s>
<s> a peck of pickled peppers peter piper picked </s>

i) Construct and show a table of probabilities for observed bigrams, including the start and end symbols, given the above corpus

Unigram frequency:

<s>	peter	piper	picked	a	peck	of	pickled	peppers	</s>
2	2	2	2	2	2	2	2	2	2

Vocabulary = V ={<s>, peter, piper, picked, a, peck, of, pickled, peppers, </s>} =|V| =10

Bi-gram frequency

	<s>	peter	piper	picked	a	peck	of	pickled	peppers	</s>
<s>	0	1	0	0	1	0	0	0	0	0
peter	0	0	2	0	0	0	0	0	0	0
piper	0	0	0	2	0	0	0	0	0	0
picked	0	0	0	0	1	0	0	0	0	1
a	0	0	0	0	0	2	0	0	0	0
peck	0	0	0	0	0	0	2	0	0	0
of	0	0	0	0	0	0	0	2	0	0
pickled	0	0	0	0	0	0	0	0	2	0
peppers	0	1	0	0	0	0	0	0	0	1
</s>	1	0	0	0	0	0	0	0	0	0

ii) Adjust and show the bigram table using Laplace smoothing (4)

	<s>	peter	piper	picked	a	peck	of	pickled	peppers	</s>
<s>	1	2	1	1	2	1	1	1	1	1

peter	1	1	3	1	1	1	1	1	1	1
piper	1	1	1	3	1	1	1	1	1	1
picked	1	1	1	1	2	1	1	1	1	2
a	1	1	1	1	1	3	1	1	1	1
peck	1	1	1	1	1	1	3	1	1	1
of	1	1	1	1	1	1	1	3	1	1
pickled	1	1	1	1	1	1	1	1	3	1
peppers	1	2	1	1	1	1	1	1	1	2
</s>	2	1	1	1	1	1	1	1	1	1

$$P_{\text{Laplace}}(w_i) = \frac{c_i + 1}{N + V}$$

Smoothed probability

	<s>	peter	Piper	picked	a	peck	of	pickled	peppers	</s>
<s>	0.083	0.166	0.083	0.083	0.166	0.083	0.083	0.083	0.083	0.083
peter	0.083	0.083	0.249	0.083	0.083	0.083	0.083	0.083	0.083	0.083
piper	0.083	0.083	0.083	0.249	0.083	0.083	0.083	0.083	0.083	0.083
picked	0.083	0.083	0.083	0.083	0.166	0.083	0.083	0.083	0.083	0.166
a	0.083	0.083	0.083	0.083	0.083	0.249	0.083	0.083	0.083	0.083
peck	0.083	0.083	0.083	0.083	0.083	0.083	0.249	0.083	0.083	0.083
of	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.249	0.083	0.083
pickled	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.249	0.083
peppers	0.083	0.166	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.166
</s>	0.166	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083

Adjusted count formula is given by

$$c^*(w_{n-1}w_n) = \frac{[C(w_{n-1}w_n) + 1] \times C(w_{n-1})}{C(w_{n-1}) + V}$$

Q13. Design a FST for consonant doubling orthographic rule. Show the example cases and exceptions if any. (3)

One-Syllable Words and Words Ending in a Single Consonant:

If a one-syllable word, or a word ending in a single consonant preceded by a single vowel, and this consonant follows a short vowel sound, you double the final consonant before adding a suffix that begins with a vowel.

Example: run + ing = running

Example: stop + ed = stopped

Example: beg + ing = begging

Words with Two or More Syllables Ending in a Single Consonant:

If a word has more than one syllable and ends in a single consonant preceded by a single vowel, and the last syllable is stressed, then you double the final consonant when adding a suffix beginning with a vowel.

Example: prefer + ed = preferred

Example: commit + ing = committing

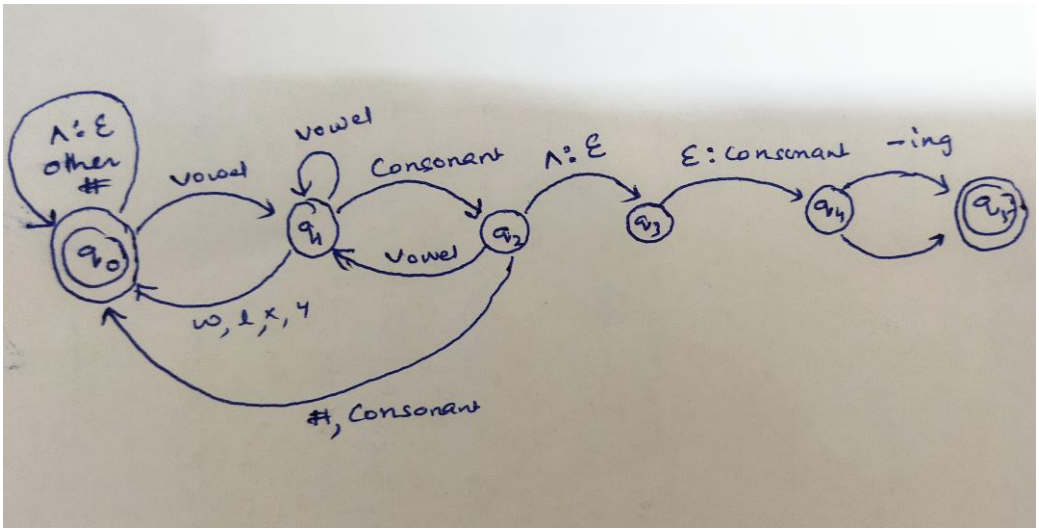
Exception

Example: equal + ity = equality (not "equalilty")

Example: model + ing = modeling (not "modelling")

Example: blow + ing = blowing (not "blowwing")

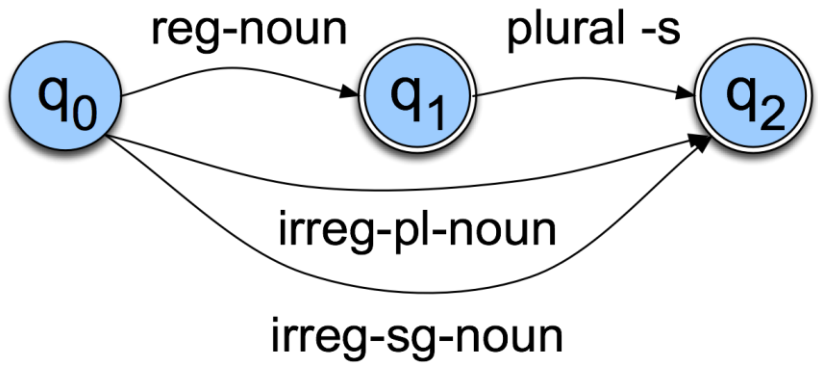
Example: follow + ed = followed (not "followed")



Q14. Write FSA for the English nominal inflection. Show the different cases in a tabular form. (3)

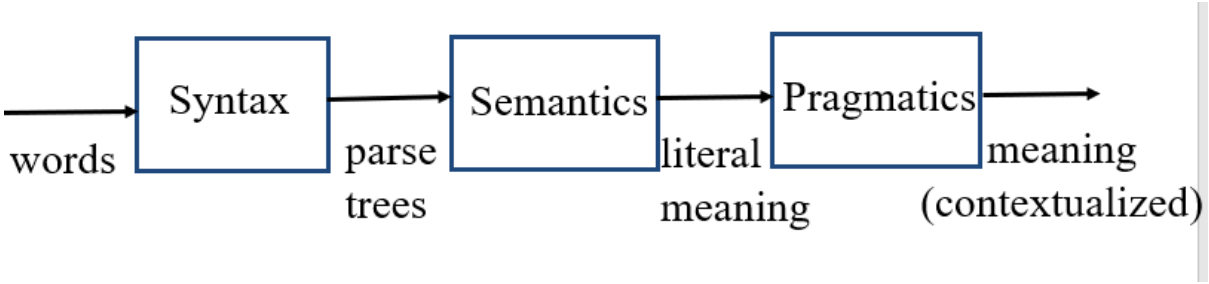
Reg-noun	Irreg-pl-noun	Irreg-sg-noun	plural

fox cat aardvark	geese sheep mice	goose sheep mouse	-s



Q15. With a neat block diagram show and explain the different categories linguistic knowledge required to process written text. (3)

1. Morphology —The study of the meaningful components of words
2. Syntax —The study of the structural relationships between words
3. Semantics — The study of meaning
4. Pragmatics — The study of how language is used to accomplish goals
5. Discourse—The study of linguistic units larger than a single utterance

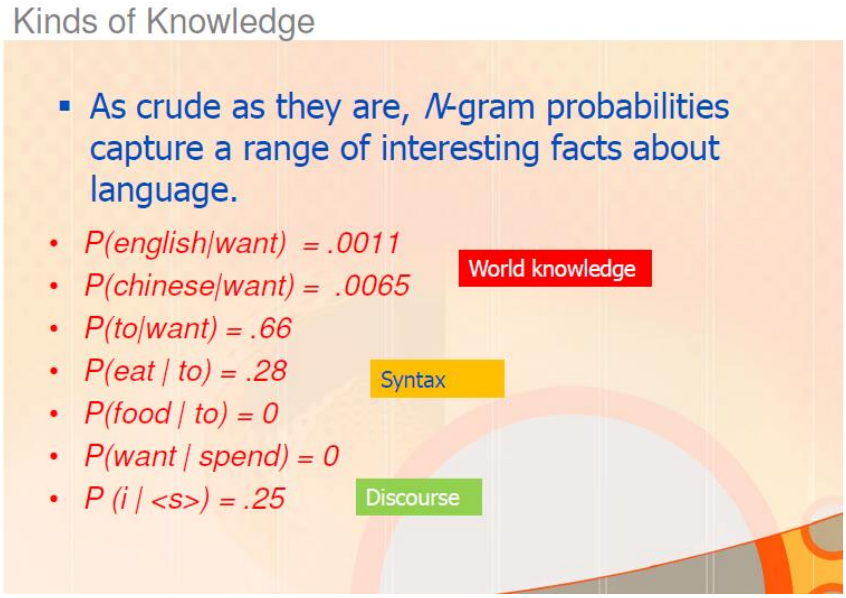


Q16. List the different types of ambiguities that occur in Natural Language Processing (3)

- Lexical (word level) ambiguity -- different meanings of words
- Syntactic ambiguity -- different ways to parse the sentence
- Interpreting partial information -- how to interpret pronouns
- Contextual information -- context of the sentence may affect the meaning of that sentence.

Q17. What are language models? How language models capture different types of knowledge. Give examples. (3)

A N-gram model is a model which computes the last word of a N-gram from the previous ones. (both N-gram and N-gram model is used in the same way). Such statistical models of word sequences are also called language models or LMs.



Q18. Assume that the following conditional probabilities are given:

P(he | <s>) = 0.09

P(wants | he) =0.21

P(to | wants) = 0.11

P(sing| to) =0.1

P(</s> | sing) =0.03

i) Write the formula you would use to compute the probability of the sentence given below using bigram model.

P(<s> he wants to sing </s>)

= P(he|<s>) P(wants | he) * P(to | wants) * P(sing| to) * P(</s> | sing)

ii) Compute the joint probability

P(<s> he wants to sing </s>)

=0.09*0.21* 0.11 * 0.1 * 0.03 = 0.000006237

(2)