CSE556: NLP (Endsem)

1. For a two-party dialogue, define the emotion recognition in conversation and the emotion-flip reasoning tasks. Also, formulate both problems appropriately. [**Hint:** Given a dialogue D = <u_1, u_2, ..., u_n> of n utterances, where u_i is an utterance of the dialogue….] **Note:** *No partial marking in formulation.* **[2+6]**

   ERC: Finding/Predicting the expressed emotion in a utterance of a dialogue/conversation.

   Given a dialogue D = <u_1, u_2, ..., u_n> of n utterances, where u_i is an utterance of the dialogue, the task of ERC is to assign an emotion $e_k \in$ E <*anger, fear, disgust, happy, sad, surprise, …*> to each utterance u_i.

   EFR: Finding/Predicting the trigger utterance(s) which is(are) responsible for the change in emotion of a speaker over a period of time in a dialogue/conversation.

   Given an emotion label e_k and a speaker $s_j \in$ S <s_1, s_2> for each utterance u_i of the dialogue D, i.e., (u_i, s_j, e_k), the task of EFR can be defined as follows: if the emotion expressed in utterance u_i changes *w.r.t.* the speaker s_j's last utterance (u_a, s_j, e_b), where $1 \leq a \leq i$, there might be a set of associated trigger-utterances u_c, where $1 \leq c \leq i$, responsible for the emotion-flip of the speaker s_j. In case of no emotion-change, we associate a 'non-trigger' label to the current utterance u_i.

2. Define sarcasm and discuss it from a humor perspective. [**Hint:** *We discussed it in the last lecture.*] **[4]**

   Sarcasm: Satirical or ironic statements usually to hurt, insult, or offend someone.

   A sarcasm is always directed towards some person, event, situation or an entity. If someone makes sarcastic comments towards someone, it will not be humorous for the target person; however, it may be humorous to the other parties or audiences.

3. Describe Claim, Fake News, Rumour, and Defamation. Give an example of an opinionated claim. **[5]**

   Claim: A claim is a statement or assertion that something is true, typically without providing evidence or proof to support it.

   Fake News: Fake news is intentionally fabricated false or misleading information presented as news.

   Rumor: Rumor is a piece of information or opinion widely disseminated with no discernible sources.

   Defamation: Defamation is the act of communicating to a third party false statements about a person, place or thing that results in damage to its reputation.

   Example of opinionated claim: The sentence "This, as I said earlier, is a complex issue." is an example of an opinionated claim.

   (Note: Opinionated claim is an assertion by a speaker who is attempting to convince others that his opinion is true. An opinion is a self-report of feelings or personal judgment and often contains clue words pointing to oneself, e.g., I think, I believe, I feel, in my opinion.)

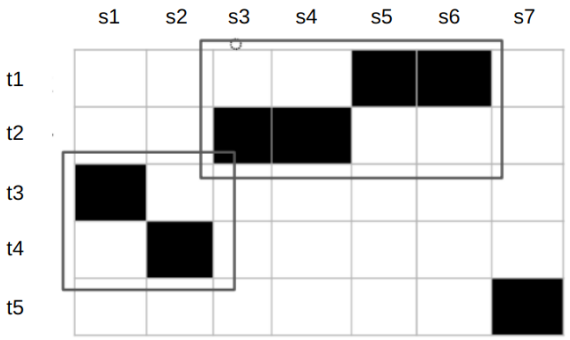
4. For aspect sentiment classification, describe two ways of instance creation with one example each. **[4]**

   Instance creation means the format of input and output that we feed to the ML/DL models, i.e., output = f(input).

   For a sentence S = <w_1, w_2, w_3, w_4, w_5, w_6, w_7, w_8, w_9, w_{10}>, where <w_3> and <w_5, w_6,> are two aspect terms (a_1 and a_2), the aspect sentiment classification instances can be designed as follows:

   - **Context-window based: assuming window of +/- 2**
     - **For a1:** <w_1, w_2, $\boldsymbol{a_1}$, w_4, w_5>
     - **For a2:** <w_3, w_4, $\boldsymbol{a_2}$, w_7, w_8>
   - **Sentence-aspect pairing**
     - **For a1:** <S, $\boldsymbol{a_1}$>
     - **For a2:** <S, $\boldsymbol{a_2}$>

5. Extract all valid phrases from the given alignment matrix. **[6]**



   t1 - s5,s6
   t2 - s3,s4
   t1,t2 - s3,s4,s5,s6
   t3 - s1
   t4 - s2
   t3,t4 - s1,s2
   t1,t2,t3,t4 - s1,s2, s3,s4,s5,s6
   t5 - s7
   t2,t3,t4 - s1,s2,s3,s4
   t1,t2,t3,t4,t5 - s1,s2, s3,s4,s5,s6,s7
   +0.6 for each correct ordering, -0.25 for any extra phrase productions.
   Num correct -> Marks
   1->1
   2->1.5
   3->2
   4->2.5
   5->3
   6->4
   7->4.5
   8->5
   9->5.5
   10 -> 6

6. In SMT, how do we compute the translation model (i.e., P(s | t))? Use equations whenever required. **[10]**

   1. Max the prob of target sentence given a source sentence. t* = $\text{argmax}_t P (t | s)$
   2. Using Bayesian rule, this can be rewritten as t* = = $\text{argmax}_t P (s | t)^* P(t)$ (As we have access to individual corpus, we can construct the p(s), p(t) n-grams as well as based on monolingual corpuses)
   3. As p(s|t) can be intractable directly, we use a latent parameter, i.e alignment model $P (s, a | t) = P(a) \prod_{i=1}^{m} P (s_i | t_{ai})$
   4. However, to know the alignment we need to know the word translation probabilities. $a^* = \text{argmax}^a \prod_{i=1}^{m} P (s_i | t_{ai})$

5. To overcome 3 and 4 we use EM algo.

7. Name three supervised paradigms of learning for NLP tasks. **[3]**

   Any NLP task can be learned in either of the three ways:
   - Classification - Sentiment, Hate-speech, FakeNews, Intent, Document Classification, etc.
   - Sequence-labeling: PoS, NER, Aspect term extraction, slot-filling, Answer extraction, etc.
   - Generation: MT, Summarization, Response generation, answer generation, explanation generation, paraphrasing, etc.

8. Write the loss function of word2vec (Mikolov et al., 2013). Assume necessary variables. **[3]**

   For skip-gram $$-\frac{1}{T}\sum_{t=1}^{T}\sum_{-c\le j\le c, j\neq 0}\log p(w_{t+j}|w_t)$$

$$\sum_{(c,w)\in D}\log\sigma(v_c v_w)+\sum_{i=1}^{k}\mathbb{E}_{w_i\sim P_n}(w_i)\left[\log\sigma(-v_i v_w)\right]$$

   -0.5 for any errors (negative sign, missing 1/T, missing summation, missing log etc.)

   Lecture 7 slide Slide 17 and 19 reference

   Similarly for CBT and negative sampling, no negative marking.

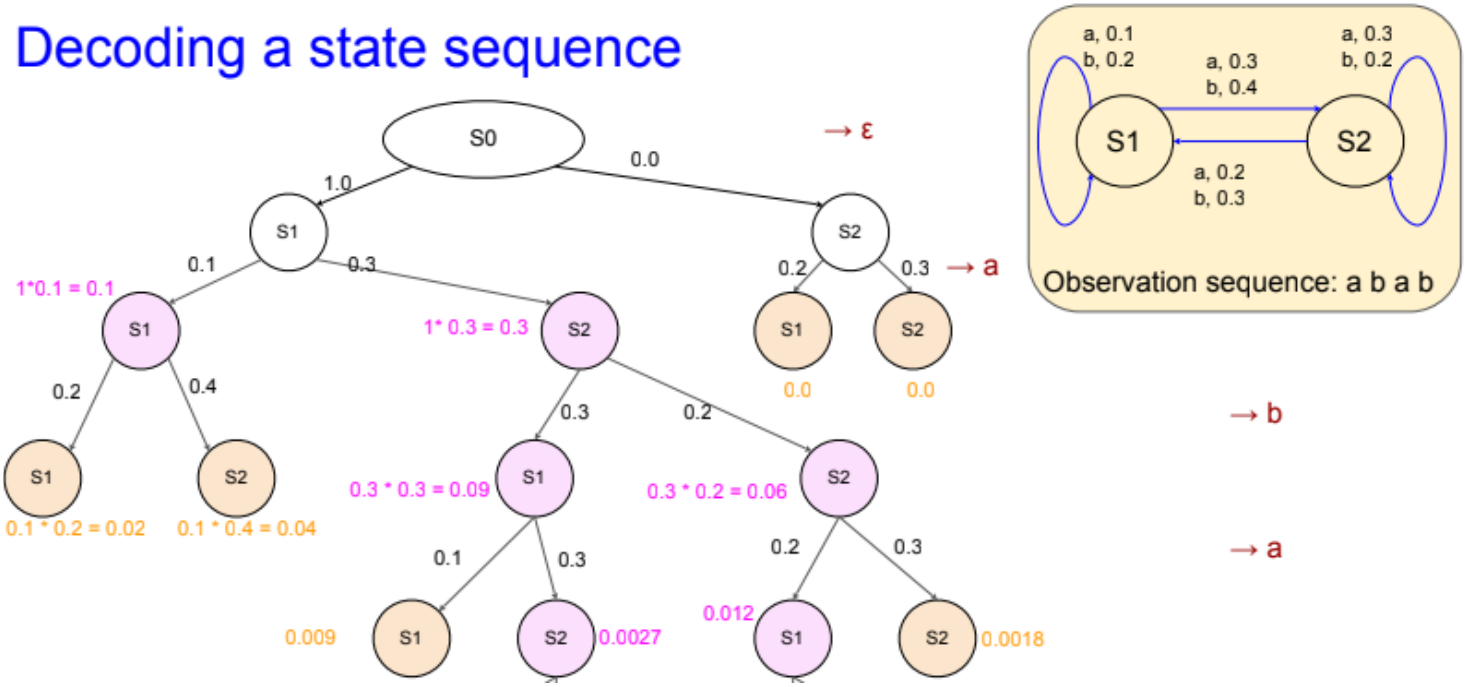
9. What do you mean by the following? Explain with examples, if possible **[1+1]**
   a. Distributed word representation: The semantic of the word is not localized to one or a few dimensions of the vector only, instead it is distributed over to all dimensions.
   b. Contextual word representation: The representation of a word differs based on the context in which it has been used, e.g., bank (finance) and bank (river) will have two distinct representations.

10. Answer any one of the following **[5]**
    a. Write the tasks of NLP hierarchy (in order) with the shortest possible definition. **OR**
       Hierarchy of the task must be maintained.
       0.5 marks for name 0.5 marks for description
       1 marks for order

Pragmatics & Discourse	*Study of semantics in context.*
Semantics	*Meaning of the sentence.*
Parsing	*Syntactic structure of the sentence.*
Chunking	*Grouping of meaningful phrases.*
Part of speech tagging	*Grammatical classes.*
Morphology	*Study of word structure.*

b. Using the Viterbi algorithm, decode the state sequence for the observation “aba”. Appropriate, probability values are depicted in the figure. Show complete decoding process. Make necessary assumptions.

Answers may differ depending on the initial probability. Here P(S1) = 1 and P(S2) = 0
1 Mark for calculating the final answer 1 Mark for Route and 3 Marks for diagram



---------- END ----------