Solution to Q2:

**STEP ONE:**

**First** we have to convert verbs and adjectives into nXn matrices, specifically three matrices

1. Adjective matrix
2. Verb Subject matrix
3. Verb Object matrix

Option 1:

Create aggregated vector(word2vec) from vocab for same POS tag, then multiply with current(word in reference) vector.

Adjectivematrix (W) = word2vec(W) \*average\_vector\_of\_all\_vocab\_POS\_of\_WT

This will not take the contextual rep into consideration.

Option 2:

As per (Mitchell and Lapata, 2010), we can use a multiplicative model to form the adjective and verb matrices where we train on small corpus to learn the tensor vector.

Another similar approach will be Baroni and Zamparelli’s (2010) lexical function model (LF) is somewhat more complex. Adjective-noun composition is modeled as the functional application of an adjective meaning (represented as a matrix) to a noun meaning (rep-  
resented as a vector).

Baroni and Zamparelli propose learning an adjective’s matrix from examples of the vectors  
for adj noun obtained directly from the corpus.  
These vectors adj noun are obtained in the same way as vectors representing a single word: when the adjective-noun combination occurs, we observe its context and construct the vector from those observations.

**STEP TWO:**

**Second step is to train the model once we have the sentence embeddings.**

We will start from a Siamese model, pass through both the sentences’s embedding through certain convolution layers then apply a logit function to train a logistic model.

Passing through convolution will dilute the embedding further.

One more option is to keep the intermediate matrices and vectors(used in stage 1) to put in the alongside the main sentence embedding, this will allow the model to learn the attention also.

References:

|  |  |
| --- | --- |
| Paper | Main author |
| Nouns are vectors, adjectives are matrices: Representing adjective-noun constructions in semantic space | Marco Baroni |
| Composition in distributional models of semantics. | Jeff Mitchell and Mirella Lapata |
| A Generalisation of Lexical Functions for Composition in Distributional Semantics | Antoine Bride |
| Learning Compositionality Functions on Word Embeddings for Modelling Attribute Meaning in Adjective-Noun Phrases | Matthias Hartung |
| A Tensor-based Factorization Model of Semantic Compositionality | Tim Van de Cruys |