Meaning and Computation: Exercise 1

Ido Shakartsi 207902834
Ofek Shapir 206480881

Question 1

1. Distance matrix is attached

2. Comparison of Similarity Measures

Word Pair	Path Sim (dist)	Resnik Sim (resnik)	Min Distance	Why Path Sim is Better
1. ('bypath', 'road')	0.50	0.0	1	"Bypath" is a direct hyponym (type) of "road" in WordNet.
2. ('byway', 'road')	0.50	0.0	1	"Byway" is highly similar to "road" structurally.
3. ('cartroad', 'road')	0.50	0.0	1	"Cartroad" is a definitional type of road.
4. ('alley', 'road')	0.25	0.0	3	"Alley" is a common type of small road/passage.
5. ('roadway', 'route')	0.50	0.0	1	"Roadway" and "route" are hypernyms/hyponyms.

The reason these pairs get Resnik sim = 0 is because their LCS is "road," which is the root of the sub-tree selected. The Resnik method is based on Information Content (IC), and the probability (P) of the root is 1. The log of 1 is 0, so the IC of the root is 0. The Resnik method thus classifies this as an uninformative measure, since it is the highest concept in the selected sub-tree.
This is why distance similarity makes more sense here, it ignores the frequency counts and just looks at the structural definitions.

3. Inter-Annotator Agreement
We agreed on all the pairs, since it really does not make sense to give a similarity measure of 0 to words who obviously have some connection to one another.

Question 2

2. The chosen word, senses and sentences:

sense = "play"
gloss1 = "a dramatic work intended for performance by actors on a stage"
gloss2 = "deliberate coordinated movement requiring dexterity and skill"

sen11 = "The critics praised the dramatic intensity of the lead actor's play at the festival last night"
sen12 = "Before the curtain rose for the evening's performance, the stage manager checked the lighting cues for the final act of the play"
sen13 = "Every actor hoped to be cast in the controversial new play written by the celebrated playwright"
sen14 = "The large, empty stage was set for the rehearsal of the experimental play that featured minimal scenery"
sen15 = "She found the performance of the new play so moving that she wept through the entire second half"
sen16 = "The grandmaster's every movement was deliberate as he prepared his next play in the end game"
sen17 = "The winning play required coordinated efforts from the entire defensive line, not just one exceptional player"
sen18 = "His incredible dexterity with the racquet allowed him to execute an untouchable drop play right on the sideline"
sen19 = "The young skater demonstrated remarkable skill in her final jump, securing a gold medal after a nearly flawless free play"
sen20 = "The coordinated fast-break play left the opposing team confused and unable to recover their defense"

2. LESK Classification Results
The LESK algorithm correctly classified 4 sentences out of 10.

Question 3

1. Word2Vec WSD Classification Results
The Word2Vec WSD algorithm correctly classified 6 out of 10 sentences.

2. Discussion of Advantages and Disadvantages

The Word2Vec WSD algorithm relies on semantic similarity in comparison to the LESK algorithm, which relies on exact word matching.

Advantages:

- **Semantic Generalization:** Word2Vec captures actual meaning, allowing it to successfully match synonyms, hypernyms, and related concepts in the context to the words in the gloss. LESK, relying on string matching, misses these semantic connections.
- **Deeper Conceptual Links:** Word2Vec vectors reflect complex relationships (like analogies) which means the similarity score captures richer conceptual proximity than simple dictionary overlap.

Disadvantages:

- **Dependence on Model Quality/Corpus:** The algorithm is heavily reliant on the training data of the pre-trained Word2Vec model. It might under-perform in domain-specific tasks or when encountering words that are out-of-vocabulary in the model, whereas the LESK algorithm relies solely on dictionary definitions.
- **Sensitivity to Noise/Equal Weighting:** Word2Vec treats all valid context words and all valid gloss words with similar importance. If there is 1 perfect match and the rest of the computed similarities are low, the perfect match will be mathematically averaged out (drowned) by the noise.
- **Computational Cost:** Word2Vec requires calculating $2n \times |L|$ cosine similarity scores (a quadratic operation), which is significantly more expensive than the simple counting required by the LESK algorithm.

3. Proposed Improvement to the Algorithm

To improve the Word2Vec WSD algorithm, we can address the fact that it assigns equal weight to all words and integrate the concept of Information Content (IC) or Inverse Document Frequency (IDF) to assign an importance measure to the gloss words.

Specifically, we would introduce a weighting factor a_x for each word x in the gloss (L), where a_x is derived from $IC(x)$ or $IDF(x)$. This ensures that a match with a rare, descriptive word in the gloss contributes far more to the final score than a match with a common, general word.